

UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO E SISTEMAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

DELINEAMENTO EXPERIMENTAL EM ENSAIOS
FATORIAIS UTILIZADOS EM PREFERÊNCIA
DECLARADA

Tese submetida à Universidade Federal de Santa Catarina para a obtenção
do título de Doutor em Engenharia de Produção.



UFSC-BU


OSMAR AMBROSIO DE SOUZA

FLORIANÓPOLIS, SC
1999

OSMAR AMBROSIO DE SOUZA

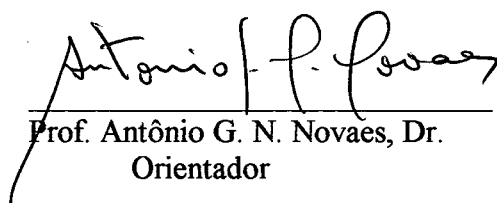
**DELINEAMENTO EXPERIMENTAL EM ENSAIOS FATORIAIS
UTILIZADOS EM PREFERÊNCIA DECLARADA**

Esta tese foi julgada adequada para a obtenção do título de **DOUTOR EM ENGENHARIA DE PRODUÇÃO** e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina.

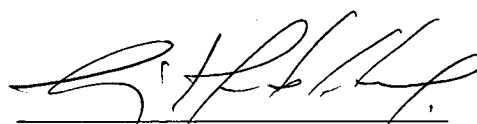


Prof. Ricardo Miranda Barcia, Ph.D.
Coordenador do Curso

Banca Examinadora:



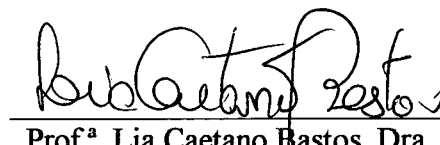
Prof. Antônio G. N. Novaes, Dr.
Orientador



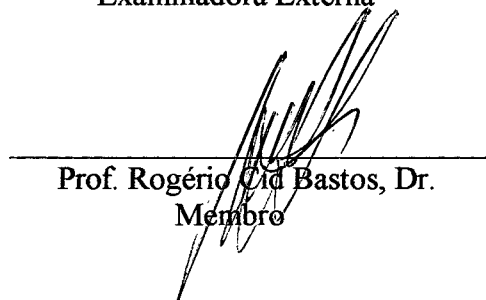
Prof. Luiz A. dos S. Senna, Ph.D.
Examinador Externo




Profª. Tiemi Matsuo, Dra.
Examinadora Externa



Profª. Lia Caetano Bastos, Dra.
Membro



Prof. Rogério Cid Bastos, Dr.
Membro



Profª. Mirian Buss Gonçalves, Dra.
Moderadora

Florianópolis, SC, 3 de agosto de 1999

À minha esposa Terezinha
e aos meus filhos Alán e Maicon.

AGRADECIMENTOS

Agradeço a Deus pela oportunidade da vida;

À minha família, pelo carinho e compreensão que me dedicaram, principalmente, nos momentos em que estive ausente para integral dedicação aos estudos;

Ao meu orientador e amigo Antônio Galvão Novaes, pelos ensinamentos, dedicação e carinho com que soube modelar minha aprendizagem. Meus sinceros agradecimentos, contigo aprendi o caminho para desvendar os mistérios da aplicação da matemática na engenharia;

Aos Professores do Departamento de Matemática da UNICENTRO, que sempre me deram apoio e incentivo durante estes anos de estudo;

À Universidade Estadual do Centro-Oeste, pelo apoio, durante a realização do curso.

Aos Professores do Departamento de Engenharia e Sistemas da UFSC, com os quais tive a oportunidade de desenvolver novas aprendizagens.

Aos Professores Membros da Banca, pela dedicação e ensinamentos transmitidos através dos comentários e sugestões, contribuíram para valorizar mais os trabalhos desenvolvidos.

À Universidade Federal de Santa Catarina: professores e funcionários, que de uma ou de outra forma contribuíram para a minha formação.

Ao meu amigo José Luiz Giraldis, pelo apoio na área de Hardware.

À Prof^a Neonila Gomes pelo apoio nos trabalhos finais da tese.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela indispensável ajuda financeira.

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE TABELAS	ix
LISTA DE ABREVEATURAS E SIGLAS	xi
RESUMO	xii
ABSTRACT	xiii

PARTE I – ASPECTOS TEÓRICOS

1 INTRODUÇÃO	1
1.1 ASSUNTO	2
1.1 FINALIDADES E OBJETIVOS	4
1.3 CONTRIBUIÇÃO CIENTÍFICA	5
1.4 ESTRUTURA DO TRABALHO	6
1.5 REVISÃO DA LITERATURA	7
2 MODELOS DE ESCOLHA DISCRETA – revisão	15
2.1 CONCEITOS	16
2.2 FUNDAMENTOS TEÓRICOS	17
2.3 MODELO <i>LOGIT MULTINOMIAL</i>	21
2.4 MODELO <i>LOGIT</i> HIERÁRQUICO	24
3 MODELO DE ESCOLHA DISCRETA MULTINOMIAL COM PROBABILIDADE CONDICIONAL	26
3.1 FORMA E COMPLEXIDADE DO EXPERIMENTO	27
3.2 TÉCNICA DE OTIMIZAÇÃO DO CONJUNTO DE ALTERNATIVAS ..	28
3.3 TÉCNICA DE OTIMIZAÇÃO DO CONJUNTO DE ESCOLHA	34
3.4 MODELO <i>LOGIT MULTINOMIAL</i> COM PROB. CONDICIONAL	39
4 AMOSTRAGEM E TESTES DE HIPÓTESES	46
4.1 VALIDAÇÃO DA AMOSTRAGEM: Testes e Procedimentos	47
4.2 VALIDAÇÃO INTERNA: Testes de Hipóteses	51
4.3 TRANSFORMAÇÃO DOS DADOS	57
4.4 IMPLEMENTAÇÃO COMPUTACIONAL – Software LMPC	62

5	METODOLOGIA PARA A ESTRUTURAÇÃO DA PESQUISA	63
5.1	FASE DA ESTRUTURAÇÃO	64
5.2	FASE DA APLICAÇÃO	73
5.3	FASE DA CONCLUSÃO	73

PARTE II – APLICAÇÃO

6	PESQUISA DE CAMPO: A utilidade dos supermercados na visão dos clientes	74
6.1	CONSIDERAÇÕES GERAIS	75
6.2	MATERIAL E MÉTODO	80
7	VALIDAÇÃO DOS DADOS AMOSTRAIS DA APLICAÇÃO	93
7.1	DADOS DESCRITIVOS DOS ATRIBUTOS SOCIAIS	94
7.2	CONDIÇÕES DA AMOSTRA – PREÇO	95
7.3	CONDIÇÕES DA AMOSTRA – QUALIDADE	97
7.4	CONDIÇÕES DA AMOSTRA – CONVENIÊNCIA	100
7.5	DADOS DE PREFERÊNCIA REVELADA	102
8	ANÁLISE E INTERPRETAÇÃO DA APLICAÇÃO	104
8.1	DADOS DE PREFERÊNCIA DECLARADA	105
8.2	DADOS DE PREFERÊNCIA REVELADA	113
8.3	VALIDAÇÃO EXTERNA DOS DADOS	114
8.4	DISCUSSÃO DOS RESULTADOS DA APLICAÇÃO	121

PARTE III – CONCLUSÃO

9	CONCLUSÕES E RECOMENDAÇÕES	126
9.1	CONCLUSÕES	127
9.2	RECOMENDAÇÕES	132

	REFERÊNCIAS BIBLIOGRÁFICAS	133
--	---	------------

LISTA DE ANEXOS

Nr.	ANEXO	Página
1	TABELA DE EXPERIMENTOS FATORIAIS	142
2	ARRANJOS ORTOGONAIS	143
3	ARRANJOS FATORIAIS	149
4	CONJUNTO DE ESCOLHA COM PROBABILIDADE CONDICIONAL ..	155
5	EXEMPLOS DE UTILIZAÇÃO DAS TABELAS	163
6	ASPECTOS COMPUTACIONAIS (Modelo LMPC)	166
7	QUESTIONÁRIO PARA ENTREVISTAS	169
8	EXEMPLOS DE CARTÕES	173
9	RELATÓRIO DAS ENTREVISTAS	177

LISTA DE FIGURAS

1.1	Modelo funil do estigma da escolha	13
2.1	Teoria econômica do consumidor	18
2.2	Esquema hierárquico com 2 níveis	25
3.1	Geometria das interações	33
3.2	Cubo	37
3.3	Relação de subconjunto para conjunto	39
4.1	Variabilidade das prob. das utilidades marginais das alternativas	59
5.1	Árvore da organização do experimento	66
6.1	Imagem da empresa	76
6.2	Organograma da aplicação	81
6.3	Mapa de avaliação geral	82
6.4	Paradigma de integração entre PR e PD	91
7.1	Dados amostrais por atributos sociais	95
7.2	Dimensionamento da amostragem : PREÇO	96
7.3	Relação entre a média e o desvio padrão – dados não transformados	98
7.4	Relação entre a média e o desvio padrão – transformação logarítmica ...	99
7.5	Dimensionamento da amostragem : QUALIDADE	99
7.6	Dimensionamento da amostragem : CONVENIÊNCIA	101
7.7	Indicação de preferências : por RENDA	103
8.1	Comparação entre as utilidades (visão geral dos clientes)	115
8.2	Comparação entre as utilidades (Classe D)	116
8.3	Comparação entre as utilidades (Classe C)	117
8.4	Comparação entre as utilidades (Classe B)	118
8.5	Comparação entre as utilidades (Classe A)	119
8.6	Coeficientes de atratividade dos supermercados	120
8.7	Relação entre comportamentos	124
-	Exemplos de cartões: PREÇO	173
-	Exemplos de cartões: QUALIDADE	174
-	Exemplos de cartões: CONVENIÊNCIA	176

LISTA DE TABELAS

1.1	Estruturação do trabalho científico	6
3.1	Efeito dos fatores e combinações	30
3.2	Delineamentos fatoriais arranjos em blocos 2^{k-p}	31
3.3	Fatorial fracionário 2^{4-1}	32
3.4	Delineamentos fatoriais fracionários 2^{k-p}	32
3.5	Arranjo ortogonal L8	33
3.6	Tabela triangular L8	33
4.1	Valor do passo	49
4.2	Valores de ϕ em função do passo A dado	50
4.3	Relação de dependência e indicação de transformação	60
5.1	Codificação para atributos com níveis discretos	69
5.2	Exemplos de codificação para atributos contínuos	69
6.1	Quadro comparativo entre os tipos de comércio	78
6.2	Classificação dos supermercados	79
6.3	Avaliação subjetiva dos atributos	79
6.4	Quadro dos atributos e os níveis : GERAL	84
6.5	Formação das alternativas	84
6.6	Quadro dos atributos e os níveis : PREÇO	85
6.7	Alternativas e níveis dos atributos : CARNES	85
6.8	Alternativas e níveis dos atributos : FRUTAS E VERDURAS	85
6.9	Alternativas e níveis dos atributos : BENS DE MERCEARIA	86
6.10	Alternativas e níveis dos atributos : LATICÍNIOS	86
6.11	Quadro dos atributos e os níveis : QUALIDADE	87
6.12	Código binários das alternativas	87
6.13	Quadro dos atributos e os níveis : CONVENIÊNCIA	88
6.14	Código binários das alternativas	88
6.15	Formação dos blocos de alternativas	89
6.16	Atributos e os níveis associados ao delineamento experimental : PR	90
6.17	Formação das alternativas	91

Continuação da lista de tabelas

7.1	Segmentação da amostra por atributos sociais	94
7.2	Distribuição das frequências dos valores da utilidade das entrevistas	97
7.3	Quadro comparativo entre os testes de Hartley e da transformação	98
7.4	Verificação do ponto de máximo da função por aproximação	101
7.5	Indicação de preferência por compras mensais	102
7.6	Indicação de preferências por tipo de supermercado	102
8.1	Estimativas dos coeficientes : GERAL	105
8.2	Estimativas dos coeficientes : PREÇO	106
8.3	Estimativas dos coeficientes : QUALIDADE	107
8.4	Estimativas dos coeficientes : CONVENIÊNCIA	108
8.5	Estimativas dos coeficientes : PREÇO (LMPC – LMN)	111
8.6	Estimativas dos coeficiente : CONVENIÊNCIA (LMPC – LMN)	112
8.7	Delineamento especial com dados PR	113
8.8	Estimativas dos coeficientes (Geral e segmentado por renda)	113
8.9	Coeficientes proporcionais estimados	114
8.10	Utilidade dos supermercados (Visão geral dos clientes)	115
8.11	Função utilidade dos supermercados (Classe D)	116
8.12	Função utilidade dos supermercados (Classe C)	117
8.13	Função utilidade dos supermercados (Classe B)	118
8.14	Função utilidade dos supermercados (Classe A)	119
8.15	Coeficiente de atratividade dos supermercados	120
8.16	Dados percentuais das utilidades	123
9.1	Problemas e soluções	129

LISTA DE ABREVIATURAS E SIGLAS

BIB	Blocos Incompletos Balanceados
BINB	Blocos Incompletos Não Balanceados
BIPB	Blocos Incompletos Parcialmente Balanceados
CAS	Coeficiente de Atratividade dos Supermercados
DA	Dimensionamento da Amostra
HM	Teste de Hausman-McFadden
IC	Intervalo de Confiança
IIA	Independência das Alternativas Irrelevantes
IID	Independentemente e Identicamente Distribuídos
Lim	Limite
LMN	<i>Logit Multinomial</i>
LMPC	<i>Logit Multinomial</i> com Probabilidade Condicional
LR	Razão de Verossimilhança
MTT	Teste de McFadden-Train-Tye
PD	Técnica de Preferência Declarada
PR	Técnica de Preferência Revelada
RC	Região de Confiança
TDD	Tratamento de Dados Discrepantes
CPC	Critério do Ponto Crítico
VUE	Valor de Utilidade da Entrevista

RESUMO

As principais dificuldades enfrentadas pelos pesquisadores, em pesquisa de preferência declarada, estão relacionadas com a estruturação do problema e essa estruturação está vinculada com a definição de uma função matemática que deverá representar o problema real.

Considerando que a representação do problema real envolve um número grande de variáveis, este trabalho apresenta dois procedimentos complementares que permitem construir delineamentos experimentais que possibilitam a inclusão de muitos atributos com um número reduzido de alternativas. O primeiro procedimento ocorre através da estruturação do experimento em áreas de interesse em que, cada área se constitui numa pesquisa de preferência declarada e, com a construção de uma função que permita relacionar as áreas entre si, gerando uma função utilidade geral que expresse o problema real. O segundo procedimento refere-se aos processos de otimização na construção do conjunto de alternativas com a incorporação dos arranjos ortogonais de Taguchi em pesquisa de preferência declarada, além de estruturar um esquema de utilização das técnicas de blocos incompletos balanceados e parcialmente balanceados na formação do conjunto de escolhas compatíveis com a capacidade humana de interpretação. Como consequência desses procedimentos, desenvolve o modelo *logit multinomial* com probabilidade condicional, que permite estimar os parâmetros com base na escolha multinomial com probabilidade condicional, preservando o pressuposto da propriedade dos eventos serem independentemente e identicamente distribuídos (IID) e apresenta 3 outros procedimentos a serem adotados para a validação dos dados amostrais: metodologia para o dimensionamento da amostra; tratamento para identificar e excluir as entrevistas consideradas discrepantes e procedimentos para testar a heterocedasticidade regular intra-alternativa e, se for o caso, utilizar a transformação logarítmica dos dados.

A validação da metodologia foi realizada através de uma pesquisa na área de marketing, com o objetivo de estabelecer a forma de relacionamento entre as variáveis latentes: preço, qualidade / organização e conveniência, utilizadas para compor uma função utilidade que representa as preferências pelos supermercados da cidade de Guarapuava, PR, para as compras mensais. A função utilidade, definida pela composição das funções advindas das áreas de interesse, mostrou ser eficaz na representação do comportamento geral dos consumidores, embora esses comportamentos sejam diferenciados quando segmentados por classes de rendas.

Os resultados obtidos através das técnicas de preferência declarada, comparados com os resultados da pesquisa de preferência revelada, expressaram uma aspiração intrínseca da população, enraizada de forma independente de suas condições sociais. No entanto, as ações realizadas em função da condição social demonstram otimizar os recursos, externando comportamentos diferenciados conforme as classes de rendas.

ABSTRACT

The main difficulties faced by researchers when applying Stated Preference (SP) techniques are related to the structure definition of the problem, and this is related with the selection of a proper mathematical function to represent the real problem.

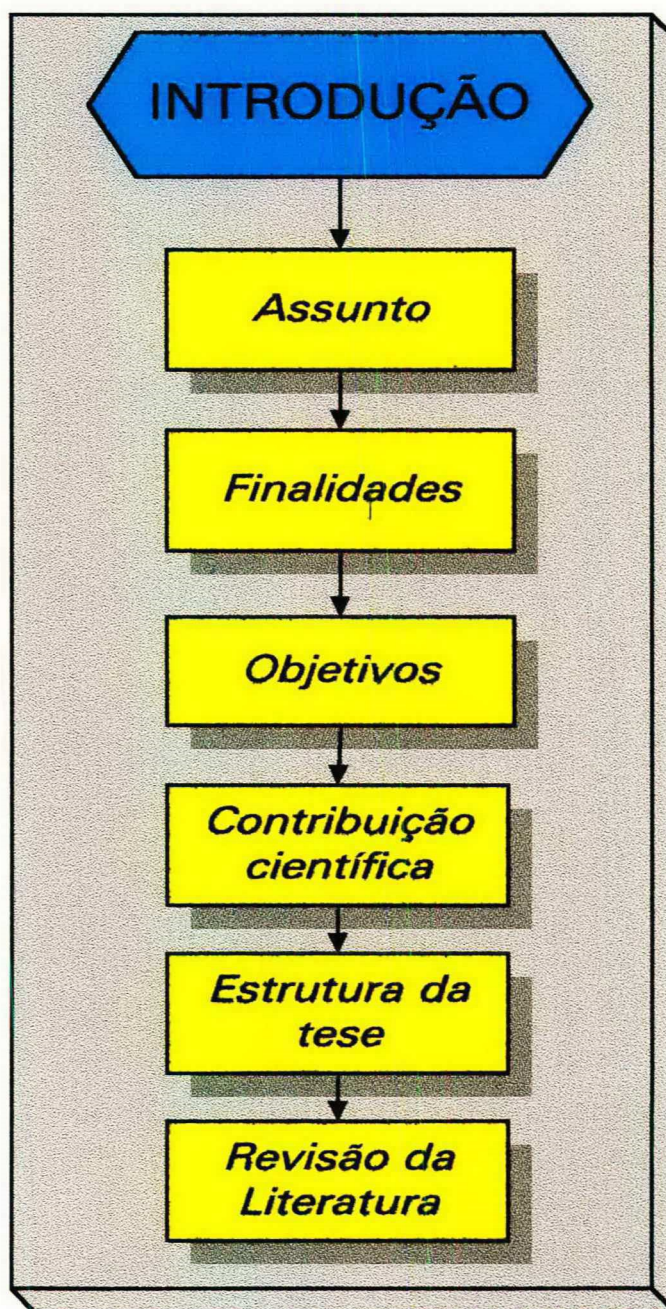
Considering that the representation of a real problem sometimes involves a great number of variables, one of the objectives of this work is to introduce two complementary procedures to SP that facilitate the inclusion of several attributes but keeping the number of alternatives low. The first procedure comprises the structuring of the experiment in separate SP surveys, each one covering a specific area of interest. It also involves the building of a function that permits the interrelation of the different areas, through a global utility function. The second procedure is related to the optimization process used to build the set of alternatives, which incorporates the orthogonal Taguchi arrangements into SP. It also comprises the structuring of incomplete balanced blocks and partially balanced blocks to be used when forming choice sets in situations where the human understanding is an important issue. As a result, we have developed a multinomial logit model with conditional probabilities, which yields parameter's estimate but respecting, at the same time, the IID property (the events being independently and identically distributed). Three other advances were also developed: (a) a methodology to define the sample size; (b) a procedure to identify and exclude outlying questionnaires, (c) a procedure to test the regular inter-alternative heterocedasticity, and, depending on the case, make the necessary data log transformation.

The validation of the methodology was done with a marketing survey intended to establish the relationship among the latent variables: price, quality/organization and convenience, which were used to compose an utility function that represents the preferences of supermarket shoppers in their monthly shopping, in the town of Gaurapuava, PR, Brazil. The utility function, defined as the weighted sum of the utilities functions obtained for each area of interest, has proved to be effective in representing the general consumer behavior, although this behavior varies when the sample is segmented according to income level.

The results obtained via SP techniques, when compared with the ones obtained through Revealed Preference methods, expressed an intrinsic aspiration of the people toward higher grade establishments, which seems to be independent of the social condition of each individual. Nevertheless, when doing their shopping in practice, they show a diverse behavior, which varies with the income level bracket.

PARTE I – ASPECTOS TEÓRICOS

Capítulo 1 - INTRODUÇÃO



1. INTRODUÇÃO

1.1. ASSUNTO

Este trabalho concentra seus estudos nas técnicas de preferência declarada (PD) e seu uso no setor de serviços, particularmente na construção do conceito para algumas variáveis latentes, atributos qualitativos de difícil mensuração, com base na combinação de vários outros atributos e níveis. Abrange, também, sobre a forma de relacionamento entre as variáveis latentes utilizadas para compor a decisão sobre o local de compras mensais, relativas aos Supermercados da Região de Guarapuava, PR, e identificar os índices de atratividade para cada tipo característico de Supermercado.

As técnicas de preferência declarada pressupõem que os indivíduos tenham condições de combinarem muitos atributos separados em alguns conceitos compostos que são usados como base para a decisão. Dessa forma, para a escolha de um supermercado, parece razoável assumir que os indivíduos possam combinar vários atributos para formar opiniões ou convicções sobre os vários locais de compras. Exemplos: “preço”, “conveniência”, “qualidade”, “variedade”, “nível de serviço”, etc..

O conceito fundamental da Técnica de Preferência Declarada aplicado em marketing surgiu da teoria de demanda do consumidor, especialmente do trabalho desenvolvido por LANCASTER (1966), quando definiu que a utilidade de um consumidor por um bem econômico pode ser decomposta em utilidades separadas referentes às características ou benefícios providas pelo bem. A visão decomposicional do processo de formação de utilidade do consumidor tornou-se amplamente aceita como uma aproximação razoável do comportamento de mercado dos consumidores, embora haja ainda debate sobre o processo envolvido na decomposição (LOUVIERE, 1994).

Segundo LOUVIERE (1994), em contraste com o fundamento conceitual, os fundamentos metodológicos estão firmemente arraigados em modelos axiomáticos ou estatísticos, embora a base particular dependa em qual dos vários paradigmas escolhe-se implementar.

Desde a sua inclusão na literatura em pesquisa de marketing, no início dos anos 70 (por exemplo, GREEN e RAO, 1971; GREEN e WIND, 1975), a análise conjunta¹ se tornou um dos métodos de pesquisa mais aplicados para entender e prever o mercado consumidor e o “trade-off”² na decisão de escolha. As técnicas de análise conjunta já eram famosas e aplicadas em psicologia (por exemplo, HOFFMAN, 1960; HOFFMAN, SLOVIC e RORER, 1968; ANDERSON, 1970), são atualmente usadas em geografia, planejamento urbano, sociologia e muitas outras áreas, e sempre tiveram destaque em pesquisa de marketing (LOUVIERE, 1994).

Um dos fatores do sucesso na pesquisa de preferência declarada está relacionado com uma boa definição da forma e da complexidade do experimento. Com relação à forma e complexidade do experimento, segundo BASTOS (1994), inicialmente, devem ser tomadas decisões com respeito a quais atributos e quantos níveis de cada um devem ser incluídos no experimento.

Quando o número de atributos e níveis é muito grande, um delineamento com fatorial completo poderia gerar muitas alternativas; sendo assim, o número de alternativas poderia ser reduzido com a adoção do delineamento com fatorial fracionário (KROES e SHELDON, 1988). Se ainda assim, segundo BASTOS (1994), o número de alternativas for muito grande, pode-se dividir em conjuntos menores. Nesse caso, um atributo comum deve ser incluído em todos os conjuntos.

Este trabalho dará ênfase à modelagem de delineamentos experimentais com a abordagem da técnica de preferência declarada utilizando os métodos de subdivisão do conjunto de alternativas, de forma que um experimento com várias alternativas e níveis possa ser avaliado em entrevistas rápidas, sem perda de realismo e com alto grau de confiabilidade; desenvolverá um modelo matemático que leva em conta a probabilidade condicional da escolha explodida, nos subconjuntos de alternativas, de forma a produzir estimativas não viesadas.

¹ Segundo KROES e SHELDON (1988), “conjoint analysis” é uma das designações utilizadas para os Métodos de obtenção de Preferência Declarada.

² Trocas compensatórias.

1.2. FINALIDADES E OBJETIVOS

A pesquisa visa abordar o assunto de forma metodológica, com desenvolvimento teórico, inédito, do modelo matemático que contempla a escolha explodida no conjunto de escolha, formado pela divisão do conjunto de alternativas, segundo as técnicas de blocos incompletos. Incorporar os arranjos ortogonais de Taguchi, arranjos fatoriais fracionários, técnicas do confundimento e quadrados latinos na constituição do conjunto de alternativas. Apresentar metodologias para tratamento da amostragem: identificação dos dados discrepantes; transformação dos dados e dimensionamento da amostra. Elaborar um plano com sugestões de ensaios fatoriais que poderão ser utilizados em outras aplicações. Desenvolver um software que possibilite a interface do usuário com a tecnologia; aplicar a metodologia em um pesquisa na área de marketing.

Objetivos:

- Desenvolver uma metodologia, a ser utilizada em técnicas de preferência declarada, que leva em conta a probabilidade condicional da ordenação no conjunto de escolha formado pela divisão do conjunto de alternativas;
- Incorporar os arranjos ortogonais de Taguchi nos delineamentos experimentais em pesquisa de preferência declarada;
- Verificar as diferentes formas de subdivisão do conjunto de alternativas em conjuntos de escolha, a serem utilizados em técnicas de preferência declarada, de forma que as estimativas dos parâmetros, teoricamente, não sejam viesadas;
- Aplicar o modelo em uma pesquisa de campo;
- Aplicar os testes estatísticos usuais ao modelo proposto;
- Desenvolver um programa para microcomputador do algoritmo proposto.

1.3. CONTRIBUIÇÃO CIENTÍFICA

O presente trabalho traz contribuição científica, para o conhecimento, nos seguintes aspectos:

- Modelo matemático da função *Logit Multinomial*, que contempla a probabilidade condicional, quando o conjunto de escolha se constitui em subconjunto do conjunto de alternativas disponíveis: Modelo *Logit Multinomial* com Probabilidade Condicional (LMPC);
- Incorporação dos arranjos ortogonais de Taguchi nos ensaios fatoriais utilizados em preferência declarada e construção de tabelas com sugestões de delineamentos experimentais;
- Metodologia para a utilização das técnicas de blocos incompletos nos processos de divisão do conjunto de alternativas em conjuntos de escolha;
- Método para a utilização da transformação dos dados, quando existe heterocedasticidade regular intra-alternativa, no modelo *Logit Multinomial*;
- Técnica de dimensionamento da amostra, com base na convergência dos erros nos parâmetros, quando a amostra tende ao infinito;
- Técnica para identificar as entrevistas discrepantes, com base na distribuição dos valores da utilidade de cada entrevista.

1.4. ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em 3 partes e 9 capítulos, de acordo com a tabela 1.1.

Tabela 1.1 – Estruturação do trabalho científico

P	Cap.	Finalidade	Seções
I	1	INTRODUÇÃO Apresenta o tema como um todo e sua posição na literatura.	1 – Assunto; 2 – Finalidades e objetivos do trabalho; 3 – Contribuição científica; 4 – Estrutura; 5 – Revisão da literatura.
	2	MODELOS DE ESCOLHA DISCRETA Apresenta a revisão, notação dos conceitos e fundamentos teórico.	1 – Conceitos; 2 – Fundamentos teóricos; 3 – Modelo <i>logit multinomial</i> 4 – Modelo <i>logit hierárquico</i> .
	3	MODELO DE ESCOLHA DISCRETA MULTINOMIAL COM PROBABILIDADE CONDICIONAL Apresenta o desenvolvimento e a fundamentação matemática do Modelo LMPC.	1 – complexidade do experimento; 2 – Otimização do conjunto de alternativas; 3 – Otimização do conjunto de escolha 4 – Modelo <i>logit multinomial</i> com probabilidade condicional (LMPC).
	4	AMOSTRAGEM E TESTES DE HIPÓTESES Apresenta os testes e procedimentos para a validação da amostra e validação interna da análise dos dados.	1 – Validação da amostragem; 2 – Validação interna; 3 – Transformação dos dados; 4 – Implementação computacional – Software LMPC.
	5	METODOLOGIA PARA A ESTRUTURAÇÃO DA PESQUISA Apresenta uma metodologia.	1 – Fase da estruturação; 2 – Fase da aplicação; 3 – Fase da conclusão.
II	6	PESQUISA DE CAMPO Apresenta o planejamento da pesquisa de campo.	1 – Considerações gerais; 2 – Material e método;
	7	VALIDAÇÃO DOS DADOS AMOSTRAIS DA APLICAÇÃO Apresenta os testes e procedimentos adotados para validar a amostragem.	1 – Atributos sociais; 2 – Dados amostrais: PREÇO; 3 – Dados amostrais: QUALIDADE; 4 – Dados amostrais: CONVENIÊNCIA. 5 – Dados de Preferência Revelada.
	8	ANÁLISE E INTERPRETAÇÃO DA APLICAÇÃO Apresenta os resultados obtidos na pesquisa.	1 – Dados de Preferência Declarada; 2 – Dados de Preferência Revelada; 3 – Integração: PD / PR. 4 – Discussão dos result. da aplicação.
III	9	CONCLUSÕES E RECOMENDAÇÕES Apresentação das conclusões e recomendações.	1 – Conclusões; e 2 – Recomendações.
-		REFERÊNCIAS BIBLIOGRÁFICAS	-
-		ANEXOS	-

1.5. REVISÃO DA LITERATURA

De acordo com KROES e SHELDON (1988), o primeiro passo para a construção de um delineamento experimental para obtenção de preferência declarada é a definição das variáveis (fatores ou atributos) de interesse e os seus valores (níveis) que serão avaliados pelos entrevistados. Um assunto associado é a especificação da forma matemática da função utilidade que expressa a hipótese do analista sobre o modo no qual os entrevistados combinam, de forma desagregada, uma avaliação global ou preferência. Como em modelagem de preferência revelada, é habitual assumir modelos compensatórios aditivos lineares da seguinte forma:

$$FU = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n. \quad \text{Onde:}$$

FU = função utilidade;

x_1 até x_n = valores dos fatores 1 até n ;

α_1 até α_n = pesos de utilidade para os fatores 1 até n .

Os fatores podem ser especificados no modelo como variáveis contínuas ou como variáveis de discretas.

Técnicas de Preferência Declarada “tradicionais” proporcionam aos entrevistados um conjunto de alternativas e lhes pedem para que expressem as preferências através da ordenação das alternativas de acordo com suas preferências, ou dando um valor de avaliação a cada alternativa. Nas experiências de escolha declarada³ (SC), é oferecida aos entrevistados uma combinação de algumas alternativas e é pedido para expressarem as escolhas indicando uma alternativa escolhida ou nomeando probabilidades escolhidas subjetivas a cada uma das alternativas.

O propósito de um delineamento experimental é definir as combinações dos níveis de todos os fatores incluídos na experiência, de tal modo que não sejam correlacionados entre as alternativas. Dado esse objetivo, o número total de alternativas pode ser definido como uma função do número de fatores e dos seus níveis. Porém, os entrevistados podem avaliar só uma quantidade bastante limitada de alternativas, entre 9 a 16 de cada vez (KROES e SHELDON, 1988); assim um delineamento que incorpora

³ Stated Choice Data (SC).

todas as possíveis combinações de todos os níveis de cada fator (fatorial completo) só pode ser usado se houver poucos fatores e níveis. MILLER (1956)⁴ admite o número 7, mais ou menos 2, como a quantidade de situações que um ser humano pode interagir.

Quando um ensaio fatorial completo gerar muitas alternativas, esse número pode ser reduzido adotando-se um projeto de ensaio fatorial fracionário, de forma que só uma seleção de todas as possíveis combinações é apresentada ao entrevistado. Se, mesmo assim, a quantidade de alternativas especificadas por um ensaio fatorial fracionário ainda é muito grande, o conjunto de alternativas pode ser dividido em subconjuntos menores, constituindo-se nos conjuntos de escolhas. Nessa hipótese, vários autores têm desenvolvido trabalhos.

Um possível procedimento é descrito em ANDERSEN et. al.(1986), em que os fatores e níveis são analisados em grupos de fatores. Com essa abordagem, um fator comum tem que ser incluído em todos os subconjuntos, com a finalidade de estabelecer uma conexão entre os fatores na determinação da utilidade global. Alternativamente pode ser utilizado um delineamento em blocos, no qual cada subconjunto de alternativas é sistematicamente apresentado a grupos diferentes de entrevistados. A estimativa da função utilidade global será o resultado de todas as entrevistas juntas, mas cada indivíduo contribuiu com a preferência de algumas das utilidades (KROES e SHELDON, 1988).

Ao adotar como variável resposta o ordenamento através de postos, GREEN e TULL (1978) sugerem que cada estímulo seja escrito em um cartão e que, inicialmente, seja pedido ao entrevistado que os agrupe em três pilhas: a primeira com os perfis que lhe agradam, a segunda com aqueles que lhe são indiferentes e a última com aqueles que lhe desagradam. Esse processo pode ser repetido até que não haja pilhas com número excessivo de cartões. Em seguida, o entrevistado deverá ordenar os cartões de cada pilha segundo sua preferência. Note-se que, devido a sua própria complexidade, esse tipo de procedimento deve ser adotado preferencialmente em entrevistas pessoais.

⁴ “These limits vary within a narrow range now commonly known as the magical number 7 plus-or-minus 2.”

A matriz “trade-off”, também conhecida na literatura como avaliação dois-por-vez, difere em muito do perfil completo. Nesse processo, os fatores serão dispostos dois a dois em tabelas e em cada uma, separadamente, as combinações de seus níveis serão ordenadas de acordo com a preferência do entrevistado (JOHNSON, 1974). Um dos inconvenientes associados a este método é a possibilidade de haver uma certa perda em termos de realismo, primeiro porque, ao se analisar apenas dois fatores por vez, estarão sendo ignoradas algumas possíveis interações envolvendo um número maior de atributos. Outra situação em que ocorre perda de realismo se dá quando existe uma relação natural forte entre dois (ou mais) fatores (correlação ambiental). Quando se tem correlação ambiental muito forte é recomendado usar o perfil completo, onde todos os atributos são avaliados simultaneamente (GREEN e SRINIVASAN, 1978) e (JOHNSON, 1974).

O argumento principal que parece favorecer o perfil completo é o fato de apresentar mais realismo nas alternativas que o procedimento de dois-por-vez. Mas é assumido que muitos atributos diminuem a habilidade para os entrevistados fazerem comparações e, assim, em muitos casos, o número de atributos é limitado entre 3 e 6. (MORIKAWA, 1989).

Segundo MORIKAWA (1989), uma vez que muitas alternativas confundem o entrevistado e prejudicam o interesse, seu uso é uma questão de troca compensatória entre a confiabilidade das estimativas e a validade dos dados coletados. Geralmente, é difícil para o entrevistado avaliar mais de 30 alternativas. MALHOTRA (1982), apud MORIKAWA (1989), investigou os efeitos do número de alternativas e o número de atributos. Ele concluiu que usando de 15 a 20 (ou 25) alternativas não se afeta o erro padrão dos parâmetros significativamente, especialmente se o número de atributos é pequeno e que a análise conjunta é um procedimento bastante robusto por avaliar preferências individuais.

GREEN (1974) discute sucintamente alguns tópicos de planejamento que têm por objetivo facilitar a fase de coleta de dados. Ele destaca que a diminuição no número de estímulos a serem apresentados depende basicamente da complexidade do modelo de preferência. Admitindo a inexistência de alguns efeitos de interação, pode-se abrir mão

da avaliação de alguns perfis através da adoção do ensaio fatorial fracionário (BOX, HUNTER e HUNTER, 1978) e (ADDELMAN, 1962).

Quando um dos fatores possui quantidade muito grande de níveis, GREEN (1974) sugere que se adote um método de estimação em múltiplos estágios. Num primeiro momento, seria pedido ao entrevistado que avaliasse apenas os níveis desse fator, por exemplo, atribuindo nota 10 ao de maior preferência, suponha nível A1, nota 0 ao de menor preferência, suponha nível A2, e aos outros, notas intermediárias, o que acabaria formando uma escala de preferência para esse atributo. A seguir, os perfis seriam construídos utilizando-se apenas os níveis A1 e A2 desse fator. O modelo seria estimado para esses dois níveis. Assim as utilidades de A1 e A2 estariam determinadas, que seriam utilizadas como limites de uma nova escala de preferência. O passo final seria obter as utilidades dos níveis intermediários; isso se dá com mudança de escala.

HENSHER (1994) cita a tabela de delineamentos experimentais apresentada em HAHN e SHAPIRO (1966) como uma indicação para o uso das técnicas de fatorial fracionários em preferência declarada, com as seguintes restrições: (i) se todos os efeitos principais forem independentes de duas ou mais interações; (ii) se o número de interações independentes permite tal fracionamento, (iii) se existem graus de liberdade suficiente, e (iv) se as alternativas geradas são compatíveis com a realidade. Cita, ainda, as experiências desenvolvidas por LOUVIERE e HENSHER (1983) e HENSHER et al (1989) a respeito do conjunto de escolha formado pelo delineamento de condicional duplo.

Com referência a fatorial fracionário é importante observar os trabalhos desenvolvidos pelos autores: KEMPTHORNE (1967), COCHRAN e COX (1978), BOX, HUNTER e HUNTER (1978) e MONTEGOMERY (1984). A literatura referente à utilização dos fatoriais fracionários em preferência declarada é bastante abrangente, no entanto, não se tem utilizado dos arranjos ortogonais de Taguchi, (TAGUCHI, 1988) e (ROSS, 1991), em delineamento experimental em técnicas de preferência declarada, embora as técnicas dos arranjos ortogonais de Taguchi apresentem “a priori” certa compatibilidade. Uma das finalidades desta tese é a de incorporar os arranjos ortogonais de Taguchi nas pesquisas de preferência declarada.

Como a quantidade de alternativas geradas pelos fatoriais fracionários ou mesmo pelos arranjos ortogonais de Taguchi ainda é superior à capacidade de interpretação do ser humano, conforme citado por MILLER (1956), há a necessidade de adotar procedimentos para reduzir a quantidade de alternativas no conjunto de escolha; assim, LOUVIERE e HENSHER (1983) e HENSHER et al (1989) apresentam, no condicional duplo, uma forma de divisão do conjunto de alternativas em conjunto de escolha com menor número de alternativa. No condicional duplo, o conjunto de alternativas é dividido em conjuntos menores com base nas técnicas de blocos incompletos balanceados, cujos principais trabalhos foram desenvolvidos por: YATES (1936), FISHER (1947), BOSE e NAIR (1939), MANN (1949), KEMPTHORNE (1967), MONTEGOMERY (1984), MEAD (1990) e outros. Já alguns trabalhos apresentaram divisão do conjunto de alternativas em conjuntos menores por processos empíricos (BASTOS, 1994; FREITAS, 1995; EFRÓN, 1995; GONÇALVES, 1995; MAY, 1996; VIEIRA, 1996).

Segundo DALY (1987), as limitações do modelo *Logit Multinomial*, no que se refere à propriedade da suposição de independência das alternativas irrelevantes sugerem, a partir de 1975, uma estrutura hierárquica para o modelo logit, denominado *Logit*⁵ “estruturado”, “hierárquico”, “seqüencial”, “aninhado”, ou “em árvore”. Esse modelo não requer o atendimento da propriedade IIA, que implica na restrição da elasticidade cruzada e ainda permite a redução do número de alternativas no conjunto de escolha pela estruturação do experimento em ramos. O modelo foi estudado por vários autores: WILLIAMS (1977), DALY e ZACHARY (1978), WESTIN e MANSKI (1979), McFADDEN (1979, 1981), GAUDRY, JARA-DIAZ e ORTÚZAR (1989).

BEN-AKIVA e BOCCARA (1995) desenvolvem um modelo de escolha discreta em dois estágios: (i) escolha em um conjunto generalizado e (ii) escolha em um dado conjunto de escolha, com o objetivo de expandir os modelos de escolha com a incorporação de uma representação probabilística na avaliação das várias alternativas.

⁵ “structured, hierarchical, sequential, nested or tree logit model.”

O processo de escolha no conjunto generalizado é definido pelo conjunto dos conjuntos de escolha e corresponde à probabilidade condicional da escolha de uma alternativa. O modelo do conjunto de escolha probabilístico é baseado na expressão sugerida por MANSKI (1977) para a probabilidade de um indivíduo n escolher a alternativa i :

$$P_n(i) = \sum_{C \in G} P_n(i/C) \cdot P_n(C) \quad \text{onde:}$$

C_n é o conjunto de alternativas apresentado ao indivíduo n ($C_n \subseteq C$), C é o conjunto universo de alternativas;

G é o conjunto de todos subconjuntos de C_n , não vazios;

$P_n(i)$ denota a probabilidade de o indivíduo n escolher a alternativa i sendo dado C_n ;

$P_n(i/C)$ é a probabilidade da escolha da alternativa i no conjunto C ;

$P_n(C)$ é a probabilidade do conjunto C_n em relação ao conjunto G .

BEN-AKIVA e BOCCARA (1995), ainda, alertam que, ao contrário do modelo *logit multinomial*, a função log-verossimilhança de um modelo em um conjunto com escolha probabilística não é globalmente convexo. Naturalmente, isso significa que o conjunto de coeficientes para os quais as derivadas parciais da log-verossimilhança são iguais a zero não necessariamente corresponde a um conjunto de coeficientes que globalmente maximizam a log-verossimilhança; desse modo, pode-se ter um ponto de máximo, mas não o ponto de máximo global. Assim, o processo de maximização da função torna-se mais complexo, dados os conhecimentos limitados que se têm a respeito do comportamento desses tipos de funções que envolvem a log-verossimilhança. Um aspecto importante da estratégia de estimação é conferir se o algoritmo converge ao mesmo ponto quando o processo começa com valores iniciais diferentes. Nesse caso, espera-se que a log-verossimilhança seja pelo menos estável na região de confiança dos coeficientes, ou seja, obtêm-se valores plausíveis. À mesma conclusão chegaram KITAMURA e LAM (1984) e SWAIT (1984) para modelos semelhantes.

Em uma visão ascendente dos conjuntos de escolha, HOROWITZ e LOUVIERE (1995) escreveram um artigo sobre o papel do conjunto de consideração na modelagem de escolha discreta. Citam ainda que, operacionalmente, o conjunto de consideração é formado pelas indicações das preferências. GENSCH e SOOFI (1995) apresentam um

artigo a respeito da estimação da informação teórica do conjunto de consideração individual. Utilizando um trabalho real na área de marketing, chegaram à conclusão de que a potência da presciência do *logit multinomial* é identificar a alternativa ineficaz (não considerada) antes de predizer a escolha da alternativa entre as escolhas no conjunto de consideração.

LAPERSONNE et al (1995) afirmam que o processo de escolha segue uma seqüência de estágios, cujo conhecimento não é novo, já fora abordado por BETTMAN (1979) em marketing, mas os recentes trabalhos têm mostrado a potência de presciência dos modelos de escolha a partir do primeiro estágio e no conjunto de consideração. SHOCHER et al (1991) apud (LAPERSONNE et al. 1995) usam uma representação gráfica para descrever os sucessivos estágios em modelo de funil (figura 1.1).

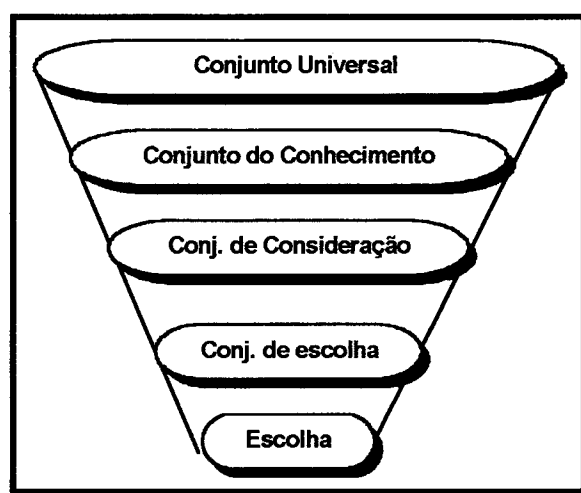


Figura 1.1 – Modelo funil do estigma da escolha

Com relação a problemas relacionados com o processo de dimensionamento da amostra, para o modelo *logit multinomial* em pesquisa de marketing, MALHOTRA (1984) cita que, para pequenas amostras ($n < 50$), os processos de estimação podem ser inadequados. As propriedades das pequenas amostras para o *logit multinomial* não foram, ainda, totalmente definidas (McFADDEN, 1974). Os procedimentos referentes ao dimensionamento têm seu valor aumentado na medida em que permitem determinar se a quantidade amostrada é suficiente para estimar assintoticamente os parâmetros, isto é, se o acréscimo de um certo número de amostra irá alterar significativamente os resultados.

Alguns autores têm escrito sobre os procedimentos para a transformação dos dados referentes aos níveis dos atributos. Por exemplo: GAUDRY e WILLS (1978), PEKELMAN e SEM (1979), ORTÚZAR e WILLUMSEN (1994) indicam a transformação de Box-Cox para a variável x e a transformação de Box-Tukey somente quando $x > 0$.

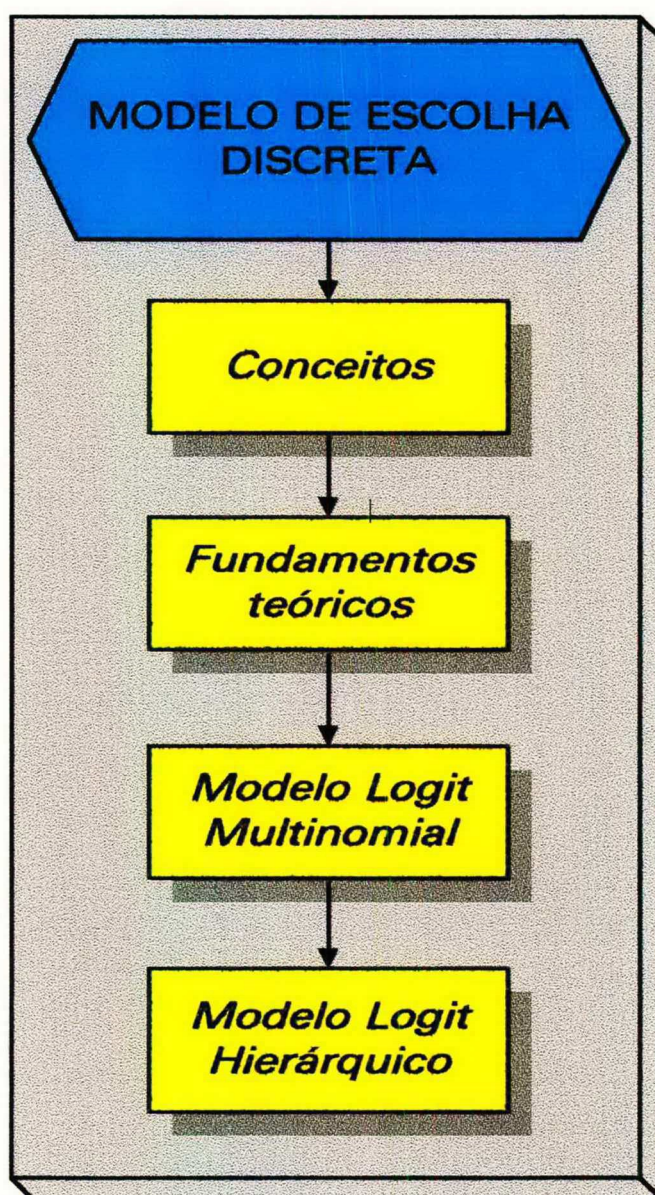
BOX, HUNTER e HUNTER (1978) apresentam uma metodologia para estabilizar a variância quando existe heterocedasticidade regular intra-alternativa, nos dados amostrais, ou seja, quando existe uma relação funcional entre a média e o desvio padrão nas escolhas efetuadas pelos entrevistados. BEN-AKIVA e LERMAN (1989) discutem, que se a suposição de homocedasticidade dos dados for inadequada, deve-se buscar uma medida satisfatória para os dados de forma a modelar corretamente a utilidade de escolha.

A proposta de pesquisa desta tese envolve a determinação de uma função utilidade pela composição das variáveis: qualidade, preço e conveniência, para modelar a utilidade dos supermercados de Guarapuava, PR, na visão de seus clientes potenciais. A construção da função utilidade, através do modelo *logit multinomial*, será confrontada com uma pesquisa de preferência revelada com o objetivo de calcular o índice de atratividade de cada tipo de supermercado. LOUVIERE & GAETH (1987) e OPPEWAL et al. (1984) desenvolveram pesquisas semelhantes, utilizando-se do Método da Integração da Informação Hierárquica (HII), respectivamente, em Supermercados e Shopping Center.

PARTE I – ASPECTOS TEÓRICOS

Capítulo 2 - MODELOS DE ESCOLHAS DISCRETAS:

Técnica de Preferência Declarada



2. MODELOS DE ESCOLHA DISCRETA

Revisão

2.1. CONCEITOS

Métodos para a obtenção de preferências declaradas referem-se a uma família de técnicas, as quais utilizam respostas individuais a respeito da preferência, em um conjunto de opções, de modo a estimar funções utilidades (KROES e SHELDON, 1988).

Para o estabelecimento do conjunto de opções, pode-se partir de descrições de situações ou contextos construídos pelo pesquisador, possibilitando, desse modo, estudar preferências que não podem ser diretamente medidas.

Para GREEN e SCRINIVASAN in SHELDON (1991), podem ser definidos como quaisquer métodos decomposicionais que estimem uma estrutura das preferências dos indivíduos utilizando sua avaliação global a respeito de um conjunto de alternativas pré-especificadas em termos de níveis de diferentes atributos.

Várias são as designações utilizadas para os métodos de obtenção de preferências declaradas. As mais conhecidas são:

- análise conjunta;
- medidas funcionais;
- análises de “trade-off” (trocas compensatórias).

Essas, utilizam procedimentos de delineamentos experimentais para gerar opções para serem avaliadas pelos entrevistados.

A abordagem de preferência declarada, embora originária da área de “marketing”, vem apresentando uma crescente aplicação em outros campos de pesquisa.

Refere-se, em essência, a um comportamento de escolha dentro de dado contexto: se você tiver essas alternativas disponíveis, qual delas você escolhe?

Entre as aplicações, pode-se citar:

- avaliação de prioridades para o desenvolvimento de várias características de um sistema público, com especial ênfase sobre fatores qualitativos;
- estimação de elasticidade de demanda para vários atributos de serviços, incluindo tarifas, frequências, etc.;
- análise de mercado e previsões;
- pesquisa e desenvolvimento de novos produtos;
- condução de estudos de planejamento para organismos governamentais.

2.2. FUNDAMENTOS TEÓRICOS

2.2.1. Teoria econômica do consumidor

O objetivo principal é atribuir significado à transformação de hipóteses sobre preferências dentro de uma função de demanda que expresse a ação de um consumidor sob dadas circunstâncias.

Um consumidor, em princípio, escolhe um produto ou serviço pelo conjunto de atributos que o produto possui, obedecendo a uma escala subjetiva de valor para cada atributo, em relação ao seu custo. Assim, o consumidor busca um produto ou serviço cujo somatório dos atributos seja igual ou superior ao valor disposto a pagar.

A base teórica comum, ou seja, o paradigma para geração de modelos da escolha discreta, é a teoria de utilidade aleatória (DOMENCICH e McFADDEN, 1975, WILLIAMS 1977) apud (ORTÚZAR e WILLUMSEN, 1994), na qual consta o postulado a respeito dos indivíduos amostrados: “Os indivíduos pertencem a uma determinada população homogênea Q, age racionalmente e possui informação, isto é, eles sempre selecionam aquela opção que maximiza sua utilidade pessoal (especificados como 'Homo economicus') sujeito legal, social, físico e orçamentário”.

De forma matemática, a teoria poderia ser representada pelo somatório:

$$\sum_{l=1}^L p_l q_l \leq I, \text{ onde } p_l \text{ representa o preço (ou peso) do atributo } q_l, l (l = 1, 2, \dots, L), \text{ e } I \text{ o}$$

valor que o consumidor está disposto a pagar. A figura 2.1 representa uma superfície de “isovalor” para I que representa todas as possíveis combinações entre 3 atributos contínuos do produto.

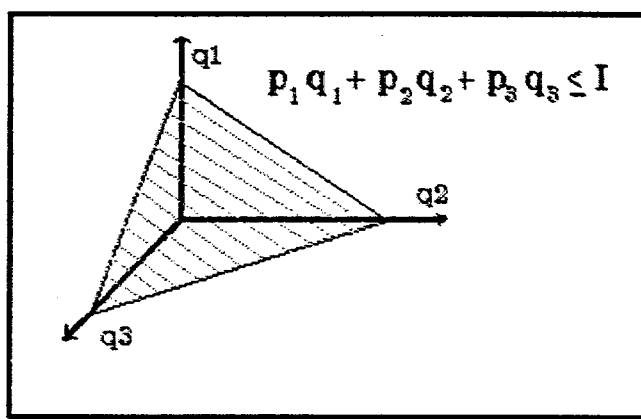


Figura 2.1 – Teoria Econômica do Consumidor.

Diante de duas alternativas (dois produtos ou serviços) $Q_i = \{q_{i1}, \dots, q_{iL}\}$ e $Q_j = \{q_{j1}, \dots, q_{jL}\}$, onde as duas alternativas apresentam diferenças nas características (níveis) dos atributos, o ato do consumidor escolher Q_i , indica que a utilidade de $Q_i \geq Q_j$.

Sob esta suposição existe uma função utilidade ordinal $U = U(q_{i1}, \dots, q_{iL})$ que expressa matematicamente a preferência do consumidor: $U(Q_i) \geq U(Q_j)$.

2.2.2. Teoria da escolha probabilística

Os primeiros trabalhos sobre a teoria da escolha probabilística foram desenvolvidos no campo da psicologia, por LUCE e SUPPES (1965). Quando um consumidor define sua preferência por uma alternativa, traz consigo uma probabilidade de escolha; no entanto, essa escolha efetuada pode não expressar sua verdadeira vontade, pode ter sido conduzido por circunstâncias momentâneas.

Assim, LUCE e SUPPES (1965) distinguem duas abordagens no mecanismo de escolha probabilística: *utilidade constante e utilidade aleatória*.

a) Utilidade Constante

Na utilidade constante, a abordagem de utilidade das alternativas é fixa. No instante em que uma alternativa é selecionada, define-se uma probabilidade de escolha da alternativa selecionada sobre o conjunto de escolha.

Denota a probabilidade de o consumidor escolher a alternativa i em um conjunto de escolha C_n por:

$$P_n(i) = P\left(\frac{i}{C_n}\right), \quad \text{onde: } 0 \leq P_n(i) \leq 1 \quad \text{e} \quad \sum P_n(i) = 1.$$

A abordagem da utilidade aleatória, formalizada por MANSKI (1977), é mais na linha da teoria do consumidor. Nessa abordagem, leva-se em conta as inconsistências inerentes ao comportamento humano. A probabilidade de escolha de uma alternativa em relação a um conjunto de escolha pode ser escrita da seguinte forma:

$$P(i/C_n) = \Pr[U_{in} \geq U_{jn}, \quad \forall j \in C_n].$$

Nessa abordagem, assume-se que a probabilidade de escolha esteja inserida em uma distribuição de probabilidade. Em geral, expressa-se como utilidade aleatória de uma alternativa a soma de dois componentes: um valor determinístico (sistemático, V_{in}) e um valor aleatório (não determinístico, ξ_{in}):

$$U_{in} = V_{in} + \xi_{in}. \quad \text{Assim: } P(i/C_n) = \Pr[V_{in} + \xi_{in} \geq V_{jn} + \xi_{jn}, \quad \forall j \in C_n].$$

- *Componente Determinístico (V):*

Assume-se que o componente determinístico (V) é uma função linear aditiva e representada pelo conjunto de atributos que influenciaram a escolha, denotada por:

$$V_j = \alpha_j + \sum_{k=1}^K \beta_{jk} X_{jk}$$

onde: $j = 1, \dots, J$ (Alternativa); $k = 1, \dots, K$ (Conjunto de atributos observados X); β = parâmetro a ser estimado que representa a contribuição de cada atributo; α = termo independente.

b) Utilidade aleatória ou componente aleatória (ξ)

Este componente aleatório refere-se a fatores não controlados no processo de escolha, tais como: erro de percepção, de entendimento, fadiga, erros de dados experimentais, etc.

Suposições a respeito do componente aleatório do modelo (Erro):

- 1) ξ (erro) é um vetor real de variáveis aleatórias e de acordo com a população amostrada, possui uma distribuição.
- 2) Essa distribuição desconhecida advém do conjunto de alternativas que devem ser identicamente distribuídas e cada um dos erros é aditivo, ou seja, são independentes.

Essas duas condições dão origem à propriedade conhecida como “*Independentemente e Identicamente Distribuída - IID*”.

No modelo Logit Multinomial, a distribuição dos erros tem uma aproximação com a distribuição de extremo valor - tipo I, ou conhecida por:

- Distribuição de dobro exponencial;
- Distribuição de Gumbel;
- Distribuição Weibull.

2.2.3. Tipos de dados para a variável dependente

Em experimentos de preferência declarada (PD) cada combinação de atributos e níveis é definida como alternativa e representa a especificação de um produto ou serviço que pode ser observado no mercado. Os dados a serem considerados, na pesquisa, na verdade se constituem numa forma de avaliação das alternativas disponíveis e podem ser métricos: avaliação (rating) ou não métricos: ordenação (ranking) e escolha (choice).

a) Ordenação das alternativas (Rank)

Ordenação das alternativas é o método mais utilizado (HENSHER, 1994). Supõe-se que um indivíduo é mais capaz de ordenar um conjunto de alternativas do que avaliar as alternativas uma a uma. No processo de ordenação é possível incorporar um

número muito grande de informações, quando se utiliza o processo de ordenação explodida, procedimento proposto por CHAPMAN e STAELIN (1982), ou seja, incorpora a ordem de escolha no processo de análise. O processo se constitui do seguinte: o entrevistado analisa cada uma das alternativas, por exemplo, um conjunto de 4 alternativas, e escolhe a preferida; exclui-se, então, essa escolhida restando 3 alternativas e, dessas, inicia-se novamente o processo de escolha, repetindo-se até restar uma alternativa.

b) Avaliação das alternativas (Rating data)

A avaliação nada mais é do que obter uma resposta métrica para cada alternativa, e é devida ao grau de preferência pela alternativa. Normalmente utiliza-se uma escala referencial de 5 ou 10 pontos. Nesse processo, também são enquadrados os estudos de comparação pareada (par-a-par), em que são estabelecidas relações semânticas de preferências.

c) Escolha das alternativas (Choice data)

Em algumas situações, onde o processo natural de escolha em que o entrevistado é conduzido a uma escolha única, há, por vezes, interesse em montar o cenário de escolha idêntico ao processo natural. Nesse caso, monta-se o conjunto de escolha das alternativas e solicita-se ao entrevistado para que escolha uma alternativa de sua preferência. Este processo assemelha-se às pesquisas de preferência revelada, com a diferença de que o pesquisador pode controlar o cenário de escolha de acordo com o objetivo da pesquisa.

2.3. MODELO LOGIT MULTINOMIAL (LMN)

O modelo Logit binário parte da suposição de que o erro $\xi_n = \xi_{jn} - \xi_{in}$ é

logisticamente distribuído, com
$$F(\varepsilon_n) = \frac{1}{1 + e^{-\mu \varepsilon_n}}, \quad \mu > 0, \quad -\infty < \xi_n < \infty.$$

A distribuição logística tende para a distribuição normal com uma boa aproximação, quando o número da amostra é grande. A suposição de que o erro ξ_n é logisticamente distribuído equivale a dizer que ξ_{jn} e ξ_{in} são independentemente e identicamente distribuídos com a distribuição Gumbel.

Sob a suposição que ξ_n é logisticamente distribuído, a escolha probabilística de uma alternativa i é dada por:

$$P_n(i) = \Pr(U_{in} \geq U_{jn})$$

$$P_n(i) = \frac{1}{1 + e^{-\mu(V_{in} - V_{jn})}} = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}} + e^{\mu V_{jn}}}$$

Para o caso da Logit Multinomial:
$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}}$$

Por conveniência, adota-se $\mu = 1$. Esse valor reflete a suposição de homocedasticidade da variância.

2.3.1. Estimativas de parâmetros

Denota N como o tamanho da amostra e define-se:

$Y_{jn} = \{1 \text{ se o entrevistado } n \text{ escolheu a alternativa } j; \text{ ZERO caso contrário; } j \text{ representa a alternativa escolhida; } C_b \text{ representa o conjunto de escolha e } V_n = \sum \beta' \chi_n.$

A função de verossimilhança para um modelo de escolha multinomial é dada por:

$$L^* = \prod_{n=1}^N \prod_{j \in C_n} P_n \left(\frac{j}{C_b} \right)^{y_{jn}} \quad (1)$$

onde para a Logit linear nos parâmetros, temos:

$$P_n \left(\frac{j}{C_b} \right) = \frac{j}{n\{C_b\}} \quad P_n \left(\frac{j}{C_b} \right) = \frac{e^{\beta' \chi_{jnb}}}{\sum_{i \in C_b} e^{\beta' \chi_{inb}}} \quad (2)$$

onde: β é o vetor de parâmetros; χ é o vetor dos valores dos níveis dos atributos.

Aplicando o logaritmo em (1) e substituindo em (2), obtem-se a função de verossimilhança para o modelo Logit Multinomial.

$$L = \sum_{n=1}^N \sum_{j \in C_b} y_{jnb} \left(\beta' X_{jnb} - \ln \sum_{i \in C_b} e^{\beta' X_{inb}} \right)$$

onde C_b é o conjunto de alternativas apresentadas aos entrevistados n .

Propriedades (Postulados):

- i) $0 \leq P(i/C) \leq 1$, para todos $i \in C$;
- ii) $\sum_{i \in C} P(i/C) = 1$;

Para o caso da função de verossimilhança em um conjunto de escolha ordenado:

$$L^* = \prod_{n=1}^N \prod_{j=1}^{J-1} \frac{e^{\beta' X_{jnb}}}{\sum_{i \in C} e^{\beta' X_{inb}}}$$

Aplicando o logaritmo na função L^* , temos a função de log-verossimilhança:

$$L = \sum_{n=1}^N \sum_{j=1}^{J-1} \left(\beta' X_{jnb} - \ln \sum_{i=j}^J e^{\beta' X_{inb}} \right)$$

Sob certas condições, McFADDEN (1974) mostra que L é globalmente côncava; assim, se existe uma solução para equação, ela é única. O estimador de máxima verossimilhança de β é consistente, assintoticamente normal e assintoticamente eficiente.

O Logit Multinomial é o mais utilizado dos modelos de escolha multinomial de dados, no entanto, existem situações que exigem a aplicação de outros tipos de modelos. Os modelos, de modo geral, são divididos em 2 classes distintas, segundo BEN-AKIVA e LERMAN (1989):

- a) generalização da logit multinomial;
 - “Random Coefficients Logit”: implementado pela Electric Power Research Institute (EPRI, 1977);

- “Ordered Logistic” – (AMEMIYA, 1975);
- “Generalized Extreme Value (GEV) Model” – (McFADDEN, 1978);

b) classe dos modelos não logit:

- “Multinomial Probit”- o conceito do “probit multinomial” aplicado à psicologia foi escrito por THURSTONE (1927);

2.4. MODELO LOGIT HIERÁRQUICO

2.4.1. Correlação e estrutura do modelo

A notação requerida para o modelo “logit” hierárquico, a partir da maximização da utilidade individual, é extremamente complicada (McFADDEN, 1981). Entretanto, para o entendimento prático do modelo, DALY (1984) propôs a especificação do modelo de probabilidade expresso em termos de probabilidade para cada alternativa.

Na literatura, segundo DALY (1987), o modelo hierárquico com 2 níveis é definido como a probabilidade, p_{ij} da escolha da alternativa i j , onde i é o nível “superior” e j o nível “inferior”:

$$p_{ij} = p_i \cdot p_{j/i} \quad p_{i/j} = \frac{e^{(V_{j/i})}}{\sum_{k \in C(i)} e^{(V_{k/i})}} \quad p_i = \frac{e^{(V_i)}}{\sum_{l \in R} e^{(V_l)}}$$

$$V_{j/i} = X_{j/i} \quad V_i = X_i + h_i \text{Log} \sum_{k \in C(i)} e^{V_{k/i}}$$

onde $C(i)$ é o conjunto das alternativas do nível inferior que partem de um nível superior da alternativa i ; R é o conjunto das alternativas do nível superior, e $X_{j/i}$ e X_i são as medidas da atratividade da alternativa j dado i , respectivamente. A teoria que conecta o modelo hierárquico com a maximização da utilidade implica que $0 < h_i \leq 1$ para todas as composições das alternativa i .

A figura 2.2 mostra um esquema de escolha hierárquica com 2 dimensões (níveis) para modelos de transporte: M (modo de transporte) e D (destinação) definido como função utilidade indireta: $U(D,M) = U^D + U^{MD}$

O componente $U(M,D)$ pode ser escrita como $U(d,m) = U_d + U_{md}$.

Com: $m = 1, \dots, M$; e $d = 1, \dots, D$.

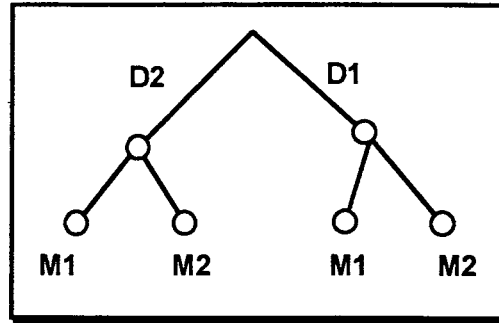


Figura 2.2 – Esquema hierárquico com 2 níveis

A utilidade das alternativas nos diferentes níveis (m, d), pode ser escrita pelos termos representados pelos componentes determinístico (V) e aleatório (e):

$$u(d,m) = V_d + V_{md} + e_d + e_{md}.$$

A matriz de covariâncias é definida por: $\sum_{(dm,d'm')} = E[e_d + e_{md}, e_{d'} + e_{d'm'}]$.

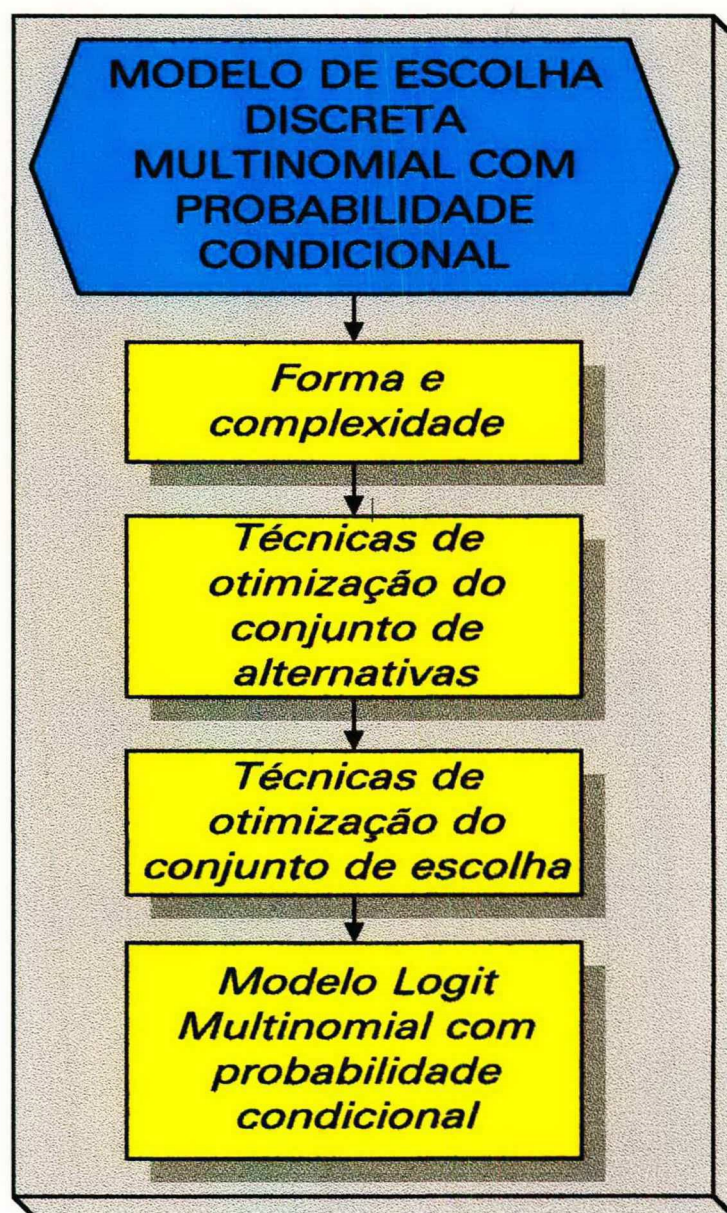
Se os componentes U_d e U_{dm} são independentemente distribuídos então:

$$E(\varepsilon_d \varepsilon_{d'}) = \sigma_D^2 \delta_{dd'}, \quad E(\varepsilon_{dm} \varepsilon_{d'm'}) = \sigma_{DM}^2 \delta_{dd'} \delta_{mm'}.$$

A matriz de variâncias: $\sum_{dm,d'm'} = \sigma_D^2 \delta_{dd'} + \sigma_{DM}^2 \delta_{dd'} \delta_{mm'}.$

PARTE I – ASPECTOS TEÓRICOS

Capítulo 3 – MODELO DE ESCOLHA DISCRETA MULTINOMIAL COM PROBABILIDADE CONDICIONAL



3. MODELO DE ESCOLHA DISCRETA MULTINOMIAL COM PROBABILIDADE CONDICIONAL

3.1. PROBLEMAS DA FORMA E COMPLEXIDADE DOS EXPERIMENTOS

Ao se adotar como variável resposta o ordenamento através de postos, GREEN e TULL (1978) sugerem que cada estímulo seja escrito em um cartão e que, inicialmente, seja pedido ao entrevistado que os agrupe em pilhas de acordo com a preferência. A matriz “trade-off” (JOHNSON, 1974) que é também conhecida na literatura como avaliação dois-por-vez difere em muito do perfil completo. Um dos inconvenientes associados a esse método “trade-off” é a possibilidade de haver uma certa perda em termos de realismo (GREEN e SRINIVASAN, 1978 e JOHNSON, 1974).

GREEN (1974) discute sucintamente alguns tópicos de planejamento que têm por objetivo facilitar a fase de coleta de dados. Ele destaca que a determinação do número de estímulos a serem apresentados depende basicamente da complexidade do modelo de preferência. Admitindo a inexistência de alguns efeitos de interação, pode-se abrir mão da avaliação de alguns perfis através da adoção de planejamento fracionário (ADDELMAN, 1962 e BOX, HUNTER e HUNTER, 1978).

Segundo HENSHER (1994), um dos aspectos mais importantes nos delineamentos estatísticos é a ortogonalidade, em que se assegura que os atributos apresentados aos indivíduos são independentemente variados uns dos outros. Essa propriedade de zero-correlação entre atributos permite ao analista empreender testes da contribuição estatística de efeitos principais e interação. Há uma visão que, embora essa seja uma propriedade desejável, não é uma condição necessária para modelagem de PD. MASON e PERREAULT (1991), apud (HENSHER, 1994), mostram que o freqüente temor dos efeitos relacionados com os atributos colineares é exagerado. Realmente o principal benefício das técnicas de preferência declarada é a habilidade para captar resposta para diversas combinações de atributos que não são observados. Esta é a razão de sua popularidade.

Nos desenvolvimentos mais recentes de experimentos de escolha explodida (KROES e SHELDON, 1988; ARTES, 1991; BASTOS, 1994; FREITAS, 1995; EFRÓN, 1995; VIEIRA, 1996; CONSTENTINO, 1997, entre outros), as respostas são feitas através da combinação das alternativas geradas pelo ensaio fatorial fracionário, em que os entrevistados são inquiridos a expressar a escolha de uma alternativa num grupo com duas ou mais alternativas, ou nomeando subjetivamente a probabilidade de escolha de cada uma das alternativas do grupo: método do perfil completo.

O propósito em questão, de um ensaio experimental, é o de definir as combinações dos níveis de todos os fatores incluídos no experimento, sem perda de realismo, de modo que eles não apresentem correlação significativa entre si, ou seja, formam delineamento de dados ortogonais (DDO).

HENSHER e BARNARD (1990) fazem uma distinção entre delineamento de dados ortogonais (DDO) e estimação de dados ortogonais (EDO), e destacam que o DDO sempre não é preservado na estimação do modelo. Considerando que a alternativa escolhida não é, “a priori”, conhecida antes de projetar o experimento, assim, não é possível projetar uma experiência que tenha DDO e que também satisfaça EDO (HENSHER e BARNARD, 1990).

3.2. TÉCNICAS DE OTIMIZAÇÃO DO CONJUNTO DE ALTERNATIVAS

Levando-se em conta que a quantidade total de alternativas é definida como uma função da quantidade de fatores e níveis envolvidos e que cada grupo de alternativas a ser submetido aos entrevistados deve ser limitado, razoavelmente, a um número pequeno de alternativas, tomando-se como referência o número de MILLER (1956), sugere-se:

- até 7 alternativas, nas pesquisas em que o entrevistado é abordado na rua ou em local que exige uma entrevista rápida;
- até 9 alternativas quando a entrevista é feita com pessoas qualificadas, com plenos conhecimentos dos assuntos abordados e realizada em local calmo e com disponibilidade de tempo.

Um modelo incorporando todas as possíveis combinações dos níveis e fatores (atributos) limitar-se-ia a poucos fatores e níveis.

Segundo HENSHER (1994), a emissão de subconjunto do conjunto de alternativas é particularmente importante quando há muitas alternativas para submeter-se ao entrevistado. Embora possa ter interesse na alternativa escolhida no subconjunto, deve-se averiguar informação adicional sobre o subconjunto pertinente. Se existem muitas repetições, é popular o uso de experiência em bloco ou aleatório, de tal modo que são pedidos aos grupos de entrevistados para responderem a um subconjunto fixo ou um subconjunto fortuito de alternativas, desde que seja assegurada a igualdade de representação nas repetições. Uma das preocupações sobre esta estratégia é a extensão de heterogeneidade de segmentação com respeito ao perfil de resposta que poderia conduzir a um distorção do perfil de resposta da população.

Foi observado por LOUVIERE (1988), apud LOUVIERE (1994) que é extremamente difícil interpretar o comportamento das interações triplas ou maiores que geralmente são excluídas. Frequentemente, a literatura cita que a maioria das variabilidades em resposta de comportamento é explicada pelo efeito principal e algumas interações duplas. Cita, ainda, que 80% ou mais da explicação são devidos aos efeitos principais e até um adicional 6% para interação de duplas. Sendo assim, quando o modelo fatorial completo for composto de muitas alternativas, esse pode ser reduzido, adotando-se:

- a técnica do confundimento ;
- o fatorial fracionário;
(KEMPTHORNE, 1967; COCHRAN e COX, 1978; BOX, HUNTER e HUNTER, 1978; MOTEGOMERY, 1984; PIMENTEL GOMES, 1987; MEAD, 1988; MCLEAN e ANDERSON, 1994)
- arranjos ortogonais desenvolvidos por Taguchi;
(TAGUCHI (1988) e ROSS (1991))

Todas as técnicas de fracionamento apresentam como fundamento básico a supressão das estimativas das interações pela junção, ou confundimento, dos efeitos principais com os efeitos das interações, que são devidas às suposições de que os atributos e níveis geram delineamentos ortogonais (DDO) e à conseqüente explicação não significativa da interação entre os efeitos.

3.2.1. Técnica do confundimento

A técnica do confundimento se constitui em uma forma de arranjar as combinações de um ensaio fatorial completo em blocos com pequenos números de combinações, utilizando-se das interações não significativas para tal fim.

Um experimento 2^k pode ser escrito como um delineamento em 2^p blocos incompletos, ou seja, 2^{k-p} , onde $p < k$. Como consequência tem-se 2, 4, 8 ou mais blocos com igual número de combinações.

Exemplo: Fatorial 2^3 confundido em 2 blocos $\Rightarrow 2^{3-1}$

- Método da combinação linear (L):

Seja: $L = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$, onde: x_i são os níveis 0 e 1 dos fatores i , neste caso, de 2 níveis; α_i são os coeficientes aplicados ao i -ésimo fator no efeito confundido (0, 1), ou (ausente e presente), ou ainda (-, +). Como se tem 2 níveis utiliza-se o módulo 2 (mod. 2), ou seja, quando a soma for igual ou superior a dois, subtrai o valor 2, para o valor de L, os resultados idênticos formam os blocos: $L = 0$ e $L = 1$.

Tabela 3.1 – Efeito dos Fatores e Combinações

Combinações	Efeito dos Fatores							
	(I)	A	B	AB	C	AC	BC	ABC
(i)	+	-	-	+	-	+	+	-
A	+	+	-	-	-	-	+	+
B	+	-	+	-	-	-	-	+
AB	+	+	+	+	-	+	-	-
C	+	-	-	+	+	-	-	+
AC	+	+	-	-	+	+		-
BC	+	-	+	-	+	-	+	-
ABC	+	+	+	+	+	+	+	+

Assim as definições de contrastes:

- (i) $000 \rightarrow L = 1.0 + 1.0 + 1.0 = 0 \pmod{2} = 0$;
- a $100 \rightarrow L = 1.1 + 1.0 + 1.0 = 1 \pmod{2} = 1$;
- b $010 \rightarrow L = 1.0 + 1.1 + 1.0 = 1 \pmod{2} = 1$;
- ab $110 \rightarrow L = 1.1 + 1.1 + 1.0 = 2 \pmod{2} = 0$;
- c $001 \rightarrow L = 1.0 + 1.0 + 1.1 = 1 \pmod{2} = 1$;
- ac $101 \rightarrow L = 1.1 + 1.0 + 1.1 = 2 \pmod{2} = 0$;
- bc $011 \rightarrow L = 1.0 + 1.1 + 1.1 = 2 \pmod{2} = 0$;
- abc $111 \rightarrow L = 1.1 + 1.1 + 1.1 = 1 \pmod{2} = 1$;

Bloco 01 : (i), ab, ac, bc; Bloco 02 : a, b, c, abc.

A identificação dos blocos nos delineamentos 2^k fatores em 2^p blocos (2^{k-p}) poderia ser feita também pelo método das combinações independentes entre os sinais das interações geradoras. Nesse caso a interação geradora seria ABC que formaria os mesmos blocos de acordo com os sinais (+) e (-).

A tabela 3.2 apresenta as interações geradoras de blocos dos fatoriais 2^{k-p} (MONTGOMERY, 1984), com interesse nos estudos de Preferência Declarada e o anexo 2 contém os esquemas detalhados para sua utilização em PD.

Tabela 3.2 – Delineamentos Fatoriais Arranjados em Blocos (2^{k-p}).

Fatores	Blocos	Tamanho	Geradores dos blocos	Interações confundidas
3	2	4	ABC	ABC
	4	2	AB, AC	AB, AC, BC
4	2	8	ABC	ABC
	4	4	ABC, ACD	ABC, ACD, BC
	8	2	AB, BC, CD	AB, BC, CD, AC, BD, AD, ABCD
5	4	8	ABC, CDE	ABC, CDE, ABDE
	8	4	ABE, BCE, CDE	ABE, BCE, CDE, AC, ABCD, BD, ADE
6	8	8	ABED, ABCD, ACE	ABEF, ABCD, ACE, BCF, BDE, CDEF, ADF
	16	4	ABF, ACF, BDF, DEF	ABE, ACF, BDF, DEF, BC, ABCD, ABDE, AD, ACDE, CE, BDF, BCEF, ABCEF, AEF, BE

3.2.2. Fatorial fracionário

Uma análise de um delineamento fatorial completo requer um número muito grande de experimentos. Como exemplo, um delineamento 2^6 , requer 64 experimentos que possuem 63 graus de liberdade (gl). No entanto, para a análise dos efeitos principais utiliza-se somente 6 gl, 15 gl para as interações duplas e 42 gl para interações maiores. Quando as interações são não significativas ou não têm sentido prático, por vezes há o interesse em eliminá-las; dessa forma é possível confundir o efeito da interação com o efeito principal de um novo fator, reduzindo o número de experimentos. Os experimentos fatoriais fracionários têm grande aplicação nas indústrias.

No caso de preferência declarada, a qual parte do princípio da ortogonalidade entre os atributos, as interações não têm efeito prático e podem ser eliminadas, ou seja, serem utilizadas para a atribuições de outros fatores principais, reduzindo o número de experimentos (combinações) a serem executados.

Um delineamento de fatorial fracionário 2^k contém 2^{k-p} experimentos (ou combinações) e são chamados como $\frac{1}{2}^p$ frações de 2^k , ou simplesmente 2^{k-p} delineamento fatorial fracionário. Esses delineamentos requerem a seleção de p geradores independentes.

Exemplo: Fatorial 2^4 gera 16 combinações, quando se utiliza o fracionário $\frac{1}{2}$, ou seja 2^{4-1} , reduzindo-se para 8 combinações pelo confundimento da interação ABC com o fator D (tabela 3.3).

Tabela 3.3 - Fatorial Fracionário 2^{4-1}

Combinações	Fatores			
	A	B	C	D = ABC
01	-	-	-	-
02	+	-	-	+
03	-	+	-	+
04	-	-	+	+
05	+	+	-	-
06	+	-	+	-
07	-	+	+	-
08	+	+	+	+

A tabela 3.4 apresenta as interações geradoras do fatorial fracionário 2^{k-p} e o anexo 2, os ensaios experimentais com interesse nos estudos de Preferência Declarada.

Tabela 3.4 – Delineamentos Fatoriais Fracionários (2^{k-p}).

Fatores	Fração	Combinações	Geradores
3	2^{3-1}	4	C = AB
4	2^{4-1}	8	D = \pm ABC
5	2^{5-1}	16	E = \pm ABCD
	2^{5-2}	8	D = \pm AB, E = \pm AC
6	2^{6-2}	16	E = \pm ABC, F = \pm BCD
	2^{6-3}	8	D = \pm AB, E = \pm AC, F = \pm BC
7	2^{7-3}	16	E = \pm ABC, F = \pm BCD, G = \pm ACD
	2^{7-4}	8	D = \pm AB, E = \pm AC, F = \pm BC, G = \pm ABC

3.2.3. Incorporando os arranjos ortogonais de Taguchi

Segundo ROSS (1991), os arranjos ortogonais se constituem numa invenção matemática cujo registro mais antigo data de 1897, por Jacques Hadamand, matemático

francês. O valor real da utilização do arranjo consiste na capacidade de avaliar diversos fatores (atributos) com um número mínimo de testes (alternativas).

São especificados, na literatura, dois tipos de arranjos ortogonais:

- arranjos com dois níveis: L4, L8, L12 e L32;
- arranjos com três níveis: L9, L18 e L27. (TAGUCHI, 1988) e (ROSS, 1991).

O número na designação do arranjo indica o número de alternativas nele contidas, por exemplo: L27 possui 27 alternativas.

As tabelas 3.5 e 3.6 apresentam, respectivamente, as configurações das alternativas e níveis de um arranjo L8 e uma tabela de interações entre as colunas do arranjo ortogonal L8.

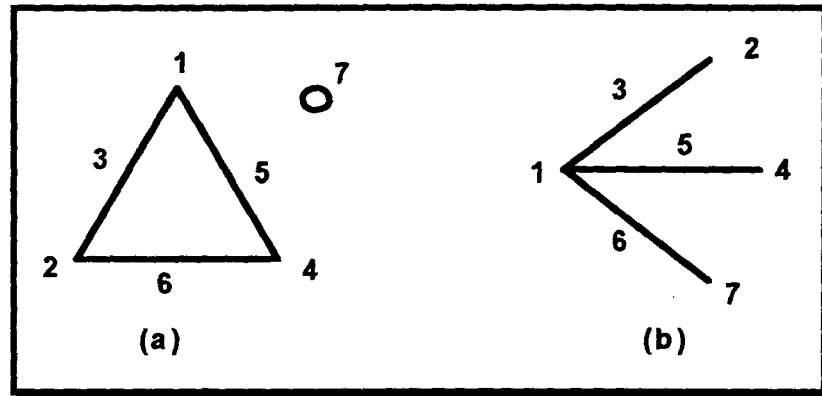
Tabela 3.5 - Arranjo Ortogonal L8

Alternativas	1	2	3	4	5	6	7
01	0	0	0	0	0	0	0
02	0	0	0	1	1	1	1
03	0	1	1	0	0	1	1
04	0	1	1	1	1	0	0
05	1	0	1	0	1	0	1
06	1	0	1	1	0	1	0
07	1	1	0	0	1	1	0
08	1	1	0	1	0	0	1

Tabela 3.6 - Tabela Triangular L8 (Interações)

Alternativas	2	3	4	5	6	7
01	3	2	5	4	7	6
02	-	1	6	7	4	5
03	-	-	7	6	5	4
04	-	-	-	1	2	3
05	-	-	-	-	3	2
06	-	-	-	-	-	1

A cada coluna de L8 pode ser designado um atributo e as designações dos atributos deve-se dar pela ordem, primeiro nas colunas correspondentes aos efeitos principais e, em seguida, nas colunas das interações, de acordo com um dos esquemas de interações na figura 3.1.



Figuras 3.1 - Geometria das Interações (L8)

Notação:	Figura (a)	Figura (b)
	Efeitos principais: 1, 2 e 4;	Efeitos principais: 1, 2, 4 e 7
	Interações : 3, 5, 6 e 7	Interações : 3, 5 e 6.

Dessa forma é possível incluir até 7 atributos com 2 níveis a um arranjo ortogonal L8. Outros arranjos possíveis com 2, 3 e 4 níveis estão disponíveis no anexo 2.

3.3. TÉCNICAS DE OTIMIZAÇÃO DO CONJUNTO DE ESCOLHA

Mesmo assim, com a utilização das técnicas de fatorial fracionário ou os arranjos ortogonais de Taguchi ainda geram, em alguns casos, um conjunto A de alternativas muito grande e o seu uso, nas técnicas de preferência declarada, estaria condicionado a uma metodologia que reduza o número de alternativas a serem apresentadas aos entrevistados sem que haja perda de realidade. Assim, deve-se estabelecer:

- subconjuntos de alternativas, A_i (ou Conjunto de escolha) que possibilite a escolha explodida, pelo entrevistado, na pesquisa de campo, conforme o número de MILLER;
- uma função que liga o subconjunto ao conjunto, fazendo com que a escolha explodida nos subconjuntos (A_i) estime uma escolha explodida no conjunto A.

Quando a quantidade de alternativas é muito grande, o material em estudo é muito heterogêneo, e existe a restrição da capacidade do ser humano em avaliar muitas

alternativas simultâneas, há necessidade de trabalhar-se com blocos incompletos, isto é, blocos que não incluam todas as alternativas, denominados:

- blocos incompletos balanceados (BIB);
- blocos incompletos parcialmente balanceados (BIPB);
- blocos incompletos não balanceados (BINB).

O anexo 4 (tabela A4.1) apresenta um resumo dos arranjos em blocos incompletos (BIB, BIPB e BINB) mais utilizados na divisão do conjunto de alternativas em conjuntos de escolhas, em que a quantidade de alternativas em cada conjunto de escolha estaria entre 3 e 6 alternativas.

3.3.1. Blocos incompletos balanceados (BIB)

Os BIB se caracterizam pelo fato de que neles, cada alternativa, aparece no mesmo bloco, com cada um das outras alternativas e sempre o mesmo número de vezes.

Notação:

- $A = \{a_1, a_2, \dots, a_v\}$ o conjunto das alternativas, onde v representa o número total de alternativas;
- $A_i = \{a_j, a_{j+1}, \dots, a_k\}$ os subconjuntos das alternativas A_i , ($i = 1, 2, \dots, b$) constituídas pelas alternativas a_j ($j = 1, 2, \dots, k$);
- r = número de vezes que uma alternativa se repete;
- v = número de alternativas do conjunto A ;
- k = número de alternativas do subconjunto A_i ;
- b = número de subconjuntos;
- λ = número de vezes que duas alternativas figuram juntas em um subconjunto A_i .

Nesses procedimentos estão envolvidas as teorias dos blocos incompletos balanceados (BIB): YATES (1936), FISHER (1938), MANN (1949), KEMPTHORNE (1967), COCHRAN e COX (1978), MONTEGOMERY (1984), MEAD (1988), PIMENTEL GOMES (1987) e outros.

Condições de balanceamentos:

- I) $b.C_{k,2} = \lambda.C_{v,2} \Rightarrow b.k.(k-1) = \lambda.v.(v-1);$
 II) $C_{k,v} = b \quad e \quad C_{(v-1),(k-1)} = r \Rightarrow b.k = r.v;$
 III) $C_{(v-2),(k-2)} = [k(k-1)].b/[v(v-1)] = \lambda. \Rightarrow b.k(k-1) = \lambda.v(v-1);$
 IV) $[k(k-1)].b/[v(v-1)] = (k-1).r/(v-1) = \lambda. \Rightarrow (k-1).r = (v-1)\lambda.;$
 MANN (1949), baseado nas Teorias de Galois, acrescenta mais que:
 V) $b \geq v$, desigualdade de FISHER (1940).

Nos experimentos onde são tomados os blocos incompletos balanceados de alternativas é muito importante identificar a eficiência (E) do experimento em relação aos

blocos completos: $E = \left(\frac{\lambda.v}{k.r} \right) = \frac{v(k-1)}{k(v-1)}$, onde $0 < E \leq 1$.

Sendo $\text{Var}(\beta) = S^2$ a estimativa da variância residual nos blocos completos, a estimativa da variância nos blocos incompletos será $\text{Var}(\beta_n) = \frac{1}{E} S^2_n = -\frac{1}{E} \mathcal{E} \left[\frac{\partial^2 L}{\partial \beta^2_n} \right]$.

Vê-se que, quanto menor o valor de E , menor a precisão das estimativas. No entanto, devemos levar em conta o seguinte: se, por um lado, reduzindo o número de alternativas em cada bloco, diminui a precisão das estimativas, por outro lado, menor número de alternativas significam menor confusão por parte do entrevistado no momento da escolha de uma alternativa. Muitas alternativas reduzem o interesse do entrevistado que passam a realizar as escolhas explodidas com menor confiabilidade (MORIKAWA, 1989), além do que existe a limitação da capacidade do ser humano de analisar simultaneamente muitas alternativas (MILLER, 1956).

3.3.2. Blocos incompletos parcialmente balanceados (BIPB)

Delineamento de blocos incompletos parcialmente balanceados são incluídos como um caso especial e, em certas circunstâncias, podem ser usados como delineamento de blocos incompletos balanceados. São caracterizados pelo fato de que as alternativas nem sempre aparecem no mesmo bloco com uma das outras alternativas o mesmo número de vezes. (BOSE e NAIR, 1939).

Vários delineamentos podem surgir por tradução de configurações geométricas. Em um poliedro regular, por exemplo, se forem considerados os pontos como alternativas e as faces como blocos, surge um delineamento de blocos incompletos parcialmente balanceados.

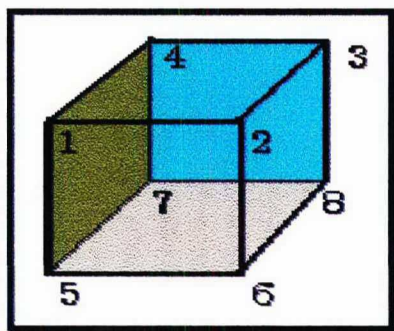


Figura 3.2 - Cubo

Por exemplo, com um cubo (figura 3.2) formam-se o arranjos:

(1, 2, 3, 4), (1, 2, 5, 6), (1, 4, 5, 7), (2, 3, 6, 8), (3, 4, 7, 8), (5, 6, 7, 8)

Nesse caso, há 3 classes de associações: os tratamentos 2, 4 e 5 fazem parte de uma classe que forma grupos com o tratamento fixo 1; os tratamentos 3, 6 e 7 fazem parte de outra classe que forma grupos com o tratamento fixo 8; e os tratamentos 1 e 8 são os tratamentos fixos. Os parâmetros do delineamento são: $t = 8$, $b = 6$, $k = 4$, $r = 3$ e Fator de eficiência $E = 14/17$.

3.3.3. Blocos incompletos não balanceados (BINB)

Os BINB são caracterizados pela divisão de um conjunto A em subconjuntos A_i , de forma que os conjuntos gerados se tornem mutuamente exclusivos. Assim, cada alternativa aparece em um dos subconjuntos, mas duas alternativas não são incluídas juntas em algum subconjunto A_i .

Notação:

- Seja $A = \{a_1, a_2, a_3, \dots, a_n\}$ o conjunto das alternativas do experimento, onde $\#A = n$ (Cardinalidade de A é n);
- Seja A_i um subconjunto do conjunto A , então $A_i = \{a_1, a_2, a_3, \dots, a_k\}$, com $i = 1, 2, 3, \dots, t$ (Subconjuntos) e $\#A_i = k$ (Cardinalidade dos subconjuntos), assim:

$$\bigcup_{i=1}^t A_i = A \quad \bigcap_{i=1}^t A_i = I \subseteq A$$

onde, nos BINB, exclusivos, $I = \emptyset$ e $k.t = \#A = n$ (cardinalidade do conjunto A).

No caso dos BIB, $I = A$ e $k.t > \#A = n$ (cardinalidade do conjunto A) e no caso dos BIPB, $I \subset A$, $I \neq \emptyset$ e $k.t > \#A = n$.

A técnica de blocos incompletos balanceados se caracteriza por apresentar as melhores condições na relação dos subconjuntos para os conjuntos.

Função de ligação: subconjunto \rightarrow conjunto

A estimativa da ordenação das alternativas no conjunto A, tomando por base a ordenação efetuada pelos entrevistados nos subconjuntos gerados através dos BIB (A_i), pode ser obtida pela aplicação do Teorema da ordenação de escolha de Luce e Suppes (LUCE e SUPPES in MORIKAWA, 1989), do seguinte modo:

$$P(C_i) = \frac{\left(\frac{C_i}{\lambda}\right)!}{v!} \Rightarrow \text{Função teórica que leva o resultado da ordenação nos}$$

subconjuntos A_i para uma ordenação dos elementos de A. Onde:

- $C_i = s_1p_1 + s_2p_2 + \dots + s_kp_k \Rightarrow$ quantidade total de alternativas que a alternativa i superou;
- p_k é a quantidade de vezes que a alternativa i é classificada no j lugar explodido, com $j = (1, 2, \dots, K)$, nos blocos A_i , tal que $\sum p_k = r$;
- s_k é a quantidade de alternativas superadas em cada bloco A_i ;
- v é o total de alternativas em A e λ o número de vezes que duas alternativas aparecem juntas.

Assim:

- $P(C_i) = [(C_i/\lambda)!] / v!$ se $C_i = \lambda m$, para qualquer $m \in \mathbb{N}$ (Conjunto dos Naturais); ou
- $P(C_i) = (C_i/\lambda)^{(C_i/\lambda)} \cdot e^{-(C_i/\lambda)} \cdot [2\pi (C_i/\lambda)]^{1/2} \cdot (v!)^{-1}$, se $C_i \neq \lambda m$ (Fórmula Stirling para valores fracionários).

A ordem de classificação da alternativa i é dada pela ordenação dos $P(C_i)$, em ordem decrescente. Observa-se, assim, que existe uma função que, aplicada aos

resultados das escolhas explodidas no BIB, gera uma ordenação perfeita nos conjuntos das alternativas originais.

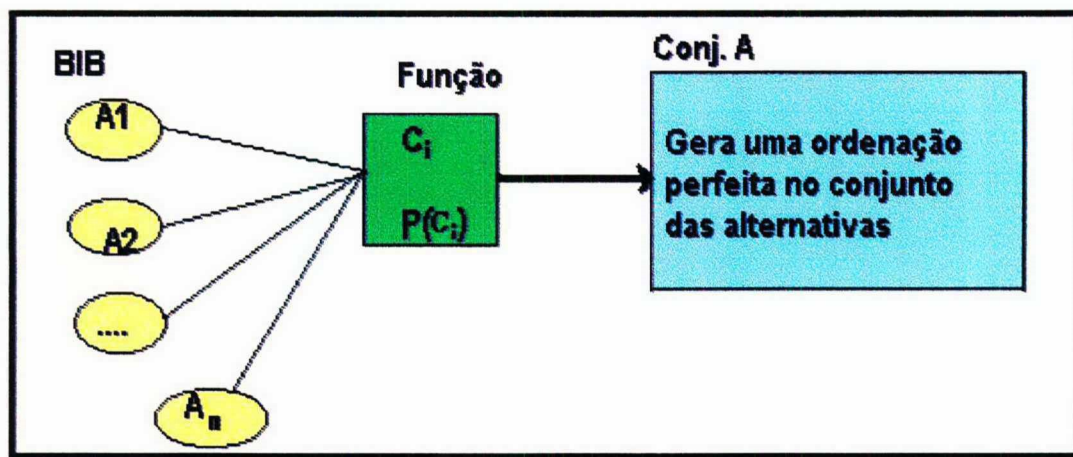


Figura 3.3 – Relação de subconjunto para conjunto

Em resumo: pode-se estimar uma ordenação de um conjunto A de alternativas, ordenando as alternativas nos subconjuntos incompletos balanceados (BIB).

Os subconjuntos A_i são os conjuntos de escolha a serem submetidos aos entrevistados para o processo de ordenação, e o conjunto A é formado de todas as alternativas advindas do arranjo ortogonal do experimento. Para os experimentos em ensaios fatoriais, com blocos incompletos, é importante levar em consideração as propriedades IID e IIA, na montagem dos modelos de estimação dos parâmetros que possuem a imposição de que os conjuntos de escolhas sejam igualmente representados na amostra.

3.4. MODELO LOGIT MULTINOMIAL COM PROBABILIDADE CONDICIONAL

O modelo matemático tem por finalidade as estimativas dos parâmetros da função utilidade (FU) oriundas da ordenação das alternativas em cada conjunto de escolha, formado pelas técnicas de BIB, submetidos aos entrevistados.

3.4.1. Fundamentos básicos

No caso em que os subconjuntos A_i forem formados pelo processo de divisão mutuamente exclusiva dos elementos de A, não se levando em conta o balanceamento, a

ordenação das alternativas numa sequência de escolha explodida (BINB), gera-se um conflito com as propriedade IID e contradição em relação ao princípio da ordenação no modelo Logit Multinomial Explodido, com o conjunto completo de escolha dividido em blocos, conforme a seguir:

- a probabilidade de uma alternativa $a_r \in A_1$ ser escolhida em primeiro lugar no subconjunto $A_1 \subset A$ é dada por:

$$P_{(a_r)} = \frac{n\{a_r\}}{n\{A_1\}} \quad \text{no caso: } \forall a_r \text{ equiprovável} \quad P_{(a_r)} = \frac{1}{k}$$

- A probabilidade de uma outra alternativa $a_s \in A_2$ ser escolhido em primeiro lugar no subconjunto $A_2 \subset A$ é dada por:

$$P_{(a_s)} = \frac{n\{a_s\}}{n\{A_2\}} \quad \text{no caso: } \forall a_s \text{ equiprovável} \quad P_{(a_s)} = \frac{1}{k}$$

E assim, a probabilidade condicional de ambas a_r e a_s , em relação ao conjunto A fica, no caso das alternativas equiprováveis:

$$P_{\left(a_r / A_1 \subset A\right)} = P_{\left(a_s / A_2 \subset A\right)} = \frac{1}{k} P_{\left(A_1 / A\right)} = \frac{1}{k} P_{\left(A_2 / A\right)} = \frac{1}{kt}$$

Logo, têm-se duas alternativas de subconjuntos distintos com a mesma probabilidade de ocorrência, dado o conjunto A , sem terem passado por um processo de comparação par a par.

O fato em questão é que temos a possibilidade de t alternativas serem classificadas em primeiro, no processo de escolha explodida das alternativas dos subconjuntos, representadas pela probabilidade condicional, sem que as mesmas tenham passado pelo processo de comparação par a par, o que vem prejudicar o princípio da boa ordem⁶. O princípio da boa ordem indica que é possível ordenar as alternativas de A , coerente com uma ordenação explodida, e uma estimativa da ordenação em A , através da ordenação dos subconjuntos A_i , neste caso, onde os subconjuntos são mutualmente

⁶ HARNSTEIN, 1970 – “LEMA DE ZORN: Seja A um conjunto não-vazio indutivamente ordenado. Então existe um elemento maximal em A .”

exclusivos, pode produzir, através do modelo Logit Multinomial Explodido, estimativas de β_A viesadas, devido a três fatores:

- Conflito com a propriedade IID;
- falta de comparação par a par em todas as alternativas de A , através dos subconjuntos A_i ;
- a influência dos pesos dos subconjuntos, ou seja, a probabilidade condicional, não é levados em conta no modelo logit multinomial explodido⁷.

FATOR 1 – Conflito com a propriedade IID

Quando os componentes da distribuição $P(y_1), P(y_2), \dots, P(y_n)$, são definidos de forma idêntica e são independentes, y_1, y_2, \dots, y_n são referidos como independentemente e identicamente distribuídos e podem ser escritos como:

$P(y_1, y_2, \dots, y_n) = P(y_1) \times P(y_2) \times \dots \times P(y_n)$, o que permite aplicar a função Logit Multinomial, $L = \prod_{i=1}^N P(y_i)$.

Essa condição só é válida se todos os eventos envolvidos são obtidos de forma idêntica, isto é, advém, no caso, de experimentos em blocos, levando-se em conta a probabilidade condicional.

Exemplo: Seja A o conjunto de alternativas e sejam A_1 e A_2 dois subconjuntos de A , onde $A_1 \cup A_2 = A$ e $A_1 \cap A_2 = \emptyset$. A probabilidade da escolha de uma alternativa i em A_1 dada por $P(i / A_1)$ e a probabilidade de outra alternativa j em A_2 dada por $P(j / A_2)$, são independentes mas não são idênticas, pois são escolhidas em cenários diferentes. Para atender à condição de idênticos, a probabilidade condicional deve ser levada em conta, assim:

$$P\left(\frac{i}{A_1} \subset A\right) = P\left(\frac{i}{A_1}\right) \times P\left(\frac{A_1}{A}\right) \quad \text{e} \quad P\left(\frac{j}{A_2} \subset A\right) = P\left(\frac{j}{A_2}\right) \times P\left(\frac{A_2}{A}\right),$$

com isso $P\left(\frac{i}{A_1} \subset A\right)$ e $P\left(\frac{j}{A_2} \subset A\right)$ são identicamente e independentemente distribuídos.

⁷ Modelo matemático citado por McFADDEN (1974), MORIKAWA (1989) e outros.

FATOR 2 - Falta da comparação par a par em todas as alternativas de A.

Dado $a_r \in A_1$, $a_s \in A_2$ e \Rightarrow

$$A_1 \cap A_2 = \emptyset \Rightarrow$$

$$P\left(\frac{a_r}{A_1 \subset A}\right) = P\left(\frac{a_s}{A_2 \subset A}\right)$$

Estabelece-se uma contradição, pois a_r não é comparado com a_s , diante do conjunto A, logo não podemos concluir em sua igualdade de probabilidade de escolha em relação ao conjunto A, já que os subconjuntos A_1 e A_2 estão mutualmente exclusivos.

FATOR 3 – As influências dos pesos dos subconjuntos.

Dado o conjunto $A = \{a_1, a_2, \dots, a_n\}$, onde n = total de alternativas e sejam os subconjuntos $A_b \subseteq A$, onde $b = 1, 2, \dots, B$ (total de subconjuntos de A com k elementos), $A_b = \{a_1, a_2, \dots, a_k\}$.

A probabilidade condicional de uma ordenação dos elementos de A_b , dado A, é

$$\begin{aligned} P\left(\frac{a_1, a_2, \dots, a_k}{A_b \subseteq A}\right) &= P\left(\frac{A_b}{A}\right) \cdot P\left(\frac{a_1, a_2, \dots, a_k}{A_b}\right) \\ &= \left[\frac{n\{A_b\}}{n\{A\}} \right] \left[\frac{n\{a_1\} \cdot n\{a_2\} \dots n\{a_k\}}{n\{A_b\}!} \right], \quad \text{onde:} \end{aligned}$$

$$\left[\frac{n\{A_b\}}{n\{A\}} \right] \Rightarrow \text{peso relativo do subconjunto } A_b \text{ no conjunto A } (\sum_b A_b) \text{ e}$$

$$\left[\frac{n\{a_1\} \cdot n\{a_2\} \dots n\{a_k\}}{n\{A_b\}!} \right] \Rightarrow \text{refere-se ao teorema da ordenação de Luce.}$$

Assim:

$$P\left(\frac{a_1, a_2, \dots, a_k}{A_b \subseteq A}\right) = \frac{\sum_{i=1}^k a_{ib}}{\sum_{b=1}^B \sum_{i=1}^k a_{ib}} \cdot \prod_{i=k}^{k-1} \frac{n\{a_i\}}{\sum_{i=j}^k a_{ib}}$$

Fruto da não observação desses fatores, as estimativas poderão ser:

$$\begin{array}{ccc} \tilde{\beta}_{A_i} & \xrightarrow{*} & \tilde{\beta}_A \approx \beta \\ (1) & & (2) \quad (3) \text{ onde:} \end{array}$$

(1) representa os coeficientes estimados através dos subgrupos A_i ;

(2) representa os coeficientes estimados através do grupo A ;

(3) representa os verdadeiros coeficientes da função;

* representa uma estimativa viesada;

\approx representa uma estimativa não-viesada.

No entanto, se os subconjuntos possuírem o mesmo peso e variabilidade e se cada subconjunto A_i for uma amostra fiel do conjunto A ou os subconjuntos forem construídos de forma balanceada (BIB), a estimativa (*) poderá ser não- viesada.

3.4.2. Função de log-verossimilhança para o Modelo LMPC

Denota N como o tamanho da amostra e define-se:

$Y_{jn} = \{1 \text{ se o entrevistado } n \text{ escolheu a alternativa } j; \text{ ZERO caso contrário.}\}$

Análogo aos procedimentos descritos no item 2.3.1, a função de verossimilhança para um modelo de escolha multinomial com probabilidade condicional é dada por:

$$L^* = \prod_{n=1}^N \prod_{j \in C_n} P_n \left(\frac{j}{C_b} \subseteq C \right)^{y_{jn}} \quad (1)$$

$C = \sum_b C_b$, j a alternativa escolhida e B = total de subconjuntos em C , onde para a Logit linear nos parâmetros, temos:

$$\begin{aligned} P_n \left(\frac{j}{C_b} \subseteq C \right) &= P_{(C_b/C)} \cdot P_{(j/C_b)} = \frac{n\{C_b\}}{n\{C\}} \times \frac{j}{n\{C_b\}} \\ P_n \left(\frac{j}{C_b} \subseteq C \right) &= \frac{\sum_{i=1}^J e^{\beta' \chi_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}}} \times \frac{e^{\beta' \chi_{jnb}}}{\sum_{i \in C_b} e^{\beta' \chi_{inb}}} \quad (2) \end{aligned}$$

Aplicando o logaritmo em (1) e substituindo em (2), obtém-se a função de log-verossimilhança para o Modelo LMPC:

$$L = \sum_{n=1}^N \sum_{j \in C_b} y_{jnb} \left(\ln \omega_b + \beta'_{jnb} - \ln \sum_{i \in C_b} e^{\beta'_{inb}} \right)$$

onde C_b é o conjunto de escolha explodido apresentado ao entrevistado n e

$$\omega_b = \frac{\sum_{i=1}^J e^{\beta'_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta'_{inb}}}$$

Propriedades (Postulados):

- i) $0 \leq P(i/C_b) \leq 1$, para todos $i \in C_b$;
- ii) $\sum_b \sum_{i \in C_b} P(C_b/C) \cdot P(i/C_b \in C) = 1$, para $\cup_b C_b = C$ e $\cap_b C_b \neq \emptyset \subseteq C$;

Para o caso da função de verossimilhança em um conjunto de escolha explodido com probabilidade condicional:

$$L^* = \prod_{n=1}^N \frac{\sum_{i=1}^J e^{\beta'_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta'_{inb}}} \times \prod_{j=1}^{J-1} \frac{e^{\beta'_{jnb}}}{\sum_{i \in C_b} e^{\beta'_{inb}}}$$

Aplicando o logaritmo na função L^* , temos a Função de Log-verossimilhança:

$$L = \sum_{n=1}^N \left[\ln \omega_b + \sum_{j=1}^{J-1} \left(\beta'_{jnb} - \ln \sum_{i=j}^J e^{\beta'_{inb}} \right) \right]$$

Todas as propriedades das estimativas de máxima verossimilhança para o modelo Logit binário se estendem para o caso multinomial com as alternativas separadas em blocos.

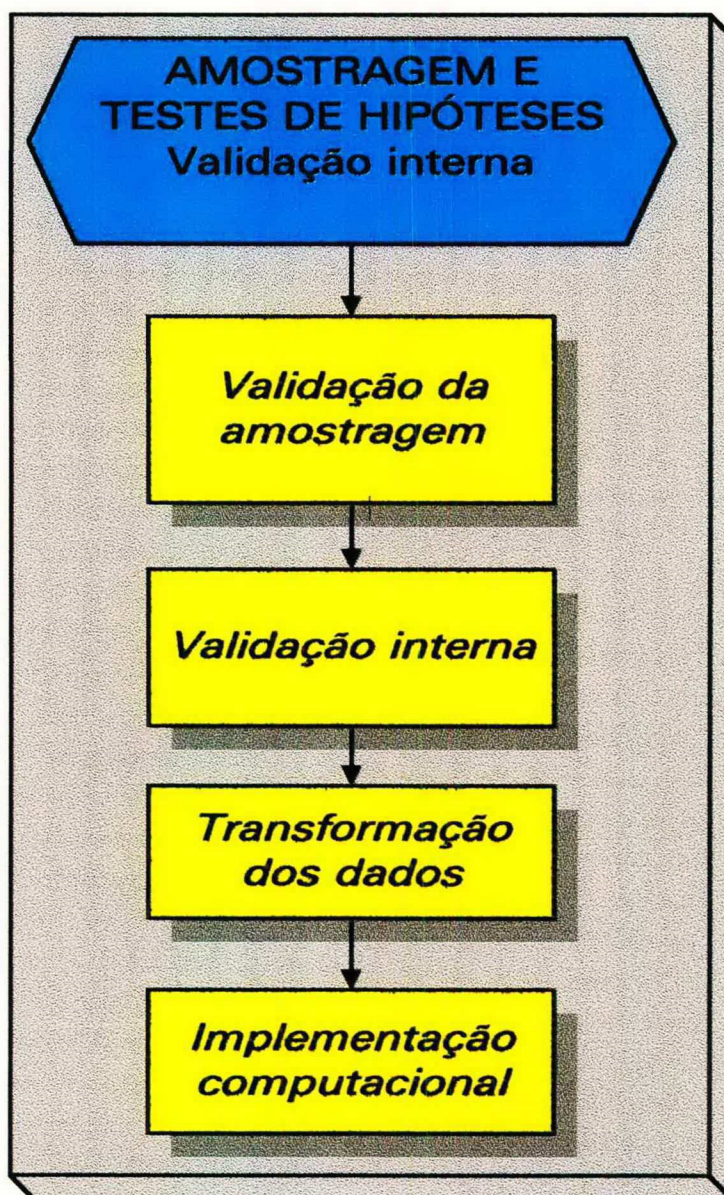
As questões levantadas sobre a influência do peso e a probabilidade condicional dos subgrupos, montados com o propósito de reduzir o número de alternativas de forma a permitir que o entrevistado possa ordená-las com certa eficiência, são levadas em conta neste modelo proposto. Os aspectos computacionais constam do anexo 6.

Segundo BEN-AKIVA e BOCCARA (1995) ao contrário do modelo logit multinomial, a função log-verossimilhança de um modelo em um conjunto com escolha probabilística não é globalmente côncava. Naturalmente, isso significa que o conjunto de coeficientes para os quais as derivadas parciais da log-verossimilhança são iguais a zero não necessariamente corresponde a um conjunto de coeficientes que globalmente maximizam a log-verossimilhança; desse modo, pode-se ter um ponto de máximo, mas não o ponto de máximo global. Assim o processo de maximização da função torna-se mais complexo, dados os conhecimentos limitados que se têm a respeito do comportamento desses tipos de funções que envolvem a log-verossimilhança. Um aspecto importante da estratégia de estimação é conferir se o algoritmo converge ao mesmo ponto quando o processo começa com valores iniciais diferentes. Nesse caso, espera-se que a log-verossimilhança seja pelo menos bem comportada na região de confiança dos coeficientes, ou seja, obtêm-se valores plausíveis. À mesma conclusão chegou KITAMURA e LAM (1984) e SWAT (1984) para modelos semelhantes.

PARTE I – ASPECTOS TEÓRICOS

Capítulo 4 - AMOSTRAGEM E TESTES DE HIPÓTESES:

Validação interna



4. AMOSTRAGEM E TESTES DE HIPÓTESES

Validação interna

4.1. VALIDAÇÃO DA AMOSTRAGEM: Testes e Procedimentos

4.1.1. Tratamento de Dados Discrepantes (TDD)

É notório ocorrer em qualquer pesquisa a inclusão de entrevistas com comportamentos que apresentem diferenças acentuadas em relação ao comportamento médio da população amostrada. Dados referentes a esses comportamentos são denominados dados discrepantes e podem ocorrer por diversas razões, no caso de PD, a mais freqüente seria a interpretação equivocada das alternativas. Mas o fato em questão é que essas entrevistas com dados discrepantes provocam resultados viesados, ou seja, resultados não afinados com a tendência natural da população amostrada.

No caso da pesquisa de preferência declarada em que o dado observado é a ordenação em um conjunto de alternativas, fica impraticável a exclusão dessas entrevistas, durante a fase da supervisão, porque não é possível 'a priori' identificar as entrevistas que se apresentem como discrepantes.

Uma solução possível para o problema em questão seria a revisão dos dados através de um algoritmo específico, que leva em conta o valor da probabilidade da utilidade individual de cada entrevista e assim determinar, por métodos descritivos, os dados que destoam do conjunto amostrado.

ALGORITMO PARA TRATAMENTO DE DADOS DISCREPANTES (TDD):

- 1) Estima-se os parâmetros (β_k^*) utilizando-se todas as entrevistas disponíveis;
- 2) De posse das estimativas dos parâmetros, β_k^* , calcula-se o valor da probabilidade da

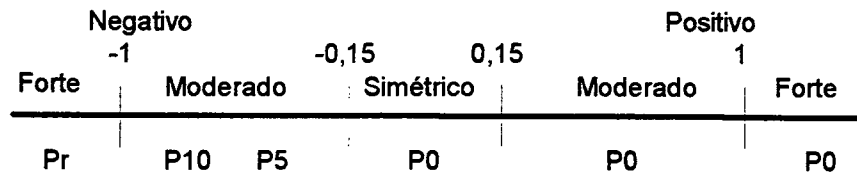
utilidade de cada entrevista, $P_n = \prod_{j=1}^{J-1} \frac{e^{\beta^* X_{ink}}}{\sum_{j \in C_n} e^{\beta^* X_{jnk}}}$, onde C_n é o conjunto

de escolha apresentado ao entrevistado n , no caso explodido, e no caso SC ('Stated

Choice'):
$$P_n = \frac{e^{\beta^* X_{ink}}}{\sum_{j \in C_n} e^{\beta^* X_{jnk}}};$$

- 3) Agrupam-se os dados, (P_n), em intervalos de classes e passa-se a calcular as estatísticas descritivas: média, variância e coeficiente de assimetria de Pearson (As);
- 4) O valor da probabilidade dos dados da entrevista será considerado discrepante se for inferior ao valor de um dado percentil de referência, (Pr). O valor do percentil é uma função do *Coeficiente de Assimetria de Pearson* (As) e é calculado em função da seguinte escala:

Escala da Indicação do Coeficiente de Assimetria de Pearson:



- a) Se $As \geq -0,15$ então não são considerados dados discrepantes;
 - b) Se $-0,5 \leq As \leq -0,15$ então adota-se o percentil 5 (P5) como valor discrepante;
 - c) Se $-1 \leq As \leq -0,5$ então adota-se o percentil 10 (P10) como valor discrepante;
 - d) Se $As \leq -1$ então $Pr = \text{menor valor entre } \{20; |\text{Inteiro}(As * 10)|\}$;
- 5) Com a identificação das entrevistas discrepantes, estas são excluídas e é feita nova estimativa dos parâmetros β_k .
 - 6) A performance, γ , da nova estimativa pode ser determinada com base nos valores da função de verossimilhança, da seguinte maneira:

$$\gamma = \left(\frac{\mathcal{L}_2(0) \cdot \mathcal{L}_1(\beta^*)}{\mathcal{L}_1(0) \cdot \mathcal{L}_2(\beta^*)} - 1 \right) \cdot 100, \text{ onde: } \gamma \text{ é dado em \% , se positivo indica melhoria de}$$

performance; $\mathcal{L}_1(0)$ e $\mathcal{L}_1(\beta^*)$ são os valores das funções de verossimilhança com os parâmetros iguais a zero e aos valores estimados, respectivamente, antes da exclusão dos dados discrepantes (índice 1) e sem os dados discrepantes (índice 2).

4.1.2. Dimensionamento da amostra (DA)

Dimensionar uma amostra é determinar o número mínimo de entrevistas necessárias para que as estimativas dos parâmetros sejam assintoticamente convergentes. E um dos métodos para se estimar esse número, em uma pesquisa de preferência declarada, consiste no seguinte:

- 1) Toma-se uma amostra inicial (piloto) com n entrevistas, com $n \geq 40$.
- 2) O valor do passo é determinado de acordo com a quantidade de entrevistas (tabela 4.1).

Tabela 4.1 – Valor do Passo

Nr de Entrevistas	Valor do Passo
De 40 a 60	3
De 61 a 100	5
Acima de 100	10

- 3) Excluir $8 \times \text{Passo}$ da quantidade n de entrevistas, resultando n_0 entrevistas, e fazem-se as estimativas de parâmetros (β_{i0});
- 4) Toma-se $n_1 = n_0 + \text{passo}$ e estimam-se novos parâmetros (β_{i1});
- 5) Calcula-se o valor do Erro: $\varphi_1 = \sqrt{\sum_i (\beta_{i1} - \beta_{i0})^2}$, que se constitui no dado explicativo para um valor n_1 . O processo se repete até o oitavo passo.
- 6) Quando o processo se repete com infinitos passos e cada repetição determina um novo φ . No limite quando $n \rightarrow \infty$ o erro $\varphi \rightarrow 0$.

Demonstração: Seja a seqüência aritmética divergente de amostras piloto, S_i :

$$S_i = n; n + A; n + 2A; n + 3A; \dots; n + (i-1)A.$$

Onde: n = amostra piloto inicial; A = quantidade iguais de entrevistas acrescentadas em cada passo i .

Quando $i \rightarrow \infty$, o limite da razão incremental $R = \frac{S_{i+1} - S_i}{S_i}$ tende a 0, ou

$$\text{seja: } \lim_{i \rightarrow \infty} \frac{S_{i+1} - S_i}{S_i} = \lim_{i \rightarrow \infty} \frac{A}{S_i} = \lim_{i \rightarrow \infty} \frac{A}{n + iA} = 0.$$

Como o valor do Erro (φ), calculado em cada passo i é uma função da razão incremental, então no limite quando $R \rightarrow 0$ os valores das estimativas β_{i+1} no passo $(i+1)$ estarão bem próximos dos valores das estimativas β_i no passo (i) tanto que o Erro $\varphi \rightarrow 0$.

- 7) Assim o dimensionamento da amostra ficará vinculado à esperança do valor de φ ser menor que um Erro predefinido (φ_n). A sequência dos valores de φ segue uma distribuição assintótica tendendo para zero.
- 8) Sendo conhecidos os pares de valores (n, φ) para os primeiros 8 termos da sequência S_i , é possível ajustar os parâmetros (a, b) da função $\varphi = F(n) = a.n^b$ e por extrapolação estimar o valor de n (dimensionamento da amostra) e IC (intervalo de confiança) tal que a estimativa de φ seja menor do que φ_n dado.
- 9) $F(n)$ pode apresentar três situações, dependendo do valor do parâmetro b :
 - a) $b > 0$: a amostra apresenta grande variabilidade e é insuficiente para determinar o dimensionamento (n) ;
 - b) $b = 0$: mostra que φ independe do número de amostras; nesse caso temos dados determinísticos;
 - c) $b < 0$: a amostra apresenta convergência e é possível estimar o dimensionamento (n) .
- 10) A tabela 4.2 apresenta os valores dos Erros φ_n a serem utilizados como base para a

determinação do dimensionamento: $n = \left(\frac{\varphi_n}{a} \right)^{\frac{1}{b}}$.

Tabela 4.2 – Valores de φ em função do passo A dado

Valor do Passo A	Valor do Erro φ_n (*)
3 entrevistas	0,01
5 entrevistas	0,02
10 entrevistas	0,05

(*) Os valores dos Erros φ_n foram estimados por processo empírico baseado em 30 experimentos de preferência declarada.

O presente processo de dimensionamento da amostra tem sua utilidade potencializada quando utilizado para verificar se a quantidade amostrada é convergente, ou seja, quando o Erro ϕ tem valor assintoticamente inferior a um valor de convergência dado, significando que qualquer acréscimo de entrevistas, necessariamente, não irá alterar significativamente os valores dos parâmetros.

4.2. VALIDAÇÃO INTERNA: Testes de Hipóteses

4.2.1. Critério do Ponto Crítico (CPC)

Tem o objetivo de verificar se os parâmetros estimados correspondem ao ponto de máximo da função para todos os pares dos atributos (X_i, X_j). Seja a função $FU = f(x_i, e)$, com $i = 1, 2, 3, \dots, I$. Toma-se x_i e x_j , com $i \neq j$ e determina-se a natureza do ponto crítico.

Suponha que $\frac{\partial f}{\partial x_i} = 0$ e $\frac{\partial f}{\partial x_j} = 0$ em P_0 .

- Se $\Delta = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 - \left(\frac{\partial^2 f}{\partial x_i^2} \right)^2 \times \left(\frac{\partial^2 f}{\partial x_j^2} \right)^2 < 0$ em P_0 , então $FU = f(x_i, x_j)$ possui:

a) Mínimo local em P_0 se $\frac{\partial^2 f}{\partial x_i^2} + \frac{\partial^2 f}{\partial x_j^2} > 0$ e

b) Máximo local em P_0 se $\frac{\partial^2 f}{\partial x_i^2} + \frac{\partial^2 f}{\partial x_j^2} < 0$.

- Se $\Delta > 0$ é Ponto de Sela e
- Se $\Delta = 0$ o Ponto é Indeterminado.

4.2.2. Teste da Razão de Verossimilhança : $LR = -2\{L(0) - L(\beta^*)\}$

Tem distribuição de χ^2 com r graus de liberdade, onde r é o número de restrições lineares (parâmetros β). Testa a hipótese de nulidade de todos os parâmetros simultaneamente. Se o valor LR for maior que o valor $\chi^2_{(\alpha, r)}$ então rejeita-se a hipótese de nulidade de todos os parâmetros simultaneamente (BEN-AKIVA e LERMAN, 1989).

4.2.3. Teste da estatística ρ^2 (Pseudo Coeficiente de Determinação) :

$$\rho^2 = 1 - \frac{L(\beta)}{L(0)}$$

É natural uma possível comparação com a estatística R^2 (Coeficiente de Determinação) que varia de 0 até 1 (ajuste perfeito). A estatística ρ^2 tem seu valor teórico também limitado de 0 a 1, mas seu valor acima de 0,2 e próximo a 0,4 indica um ajuste considerado excelente para o modelo logit multinomial (ORTUZAR e WILLUMSEN, 1990).

ρ^2_{Aj} = valor de ρ^2 ajustado para r graus de liberdade, é definido como:

$$\rho^2_{Aj} = 1 - \frac{L(\beta) - r}{L(0)}.$$

4.2.4. Teste t e Intervalo de Confiança

O teste t para a significância de um parâmetro β_k :

- Para conjunto de escolha completo (BEN-AKIVA e LERMAN (1989):

$$Var(\beta_k) = S_k^2 = -\mathcal{E} \left[\frac{\partial^2 L}{\partial \beta_k^2} \right] \quad \text{e} \quad t = \frac{\beta_k}{S_k}.$$

IC(β_k) : $\beta_k \pm t_\alpha S_k$, onde t_α é o valor crítico tabelado, na prática, toma-se $t = 2$ para $n \geq 60$ e indica, com aproximadamente 95% de probabilidade, que o β_k^* verdadeiro está no intervalo.

- Para conjuntos de escolha formado por blocos incompletos:

$$Var(\beta_k) = \frac{1}{E} S_k^2 = -\frac{1}{E} \mathcal{E} \left[\frac{\partial^2 L}{\partial \beta_k^2} \right] \quad \text{e} \quad t = \frac{\beta_k \sqrt{E}}{S_k}, \text{ onde } E = \text{Eficiência..}$$

$$IC(\beta_k) : \beta_k \pm t_\alpha \frac{S_k}{\sqrt{E}}, \text{ onde } t_\alpha \text{ é o valor crítico tabelado.}$$

4.2.5. Teste de comparação entre as utilidades das alternativas

Considere as variáveis aleatória y_1, y_2, \dots, y_k , não necessariamente normais, com as médias definidas pelos componentes do vetor β (Vetor de parâmetros) e a matriz de

variâncias definida por $-\mathcal{E} \left[\frac{\partial^2 L}{\partial \beta_k^2} \right]$.

A média e a variância de uma combinação linear $Y_v = a_1y_1 + a_2y_2 + \dots + a_ky_k$, com a_i valores constantes são:

$$\mathcal{E}(Y_v) = a_1\beta_1 + a_2\beta_2 + \dots + a_k\beta_k \quad e$$

$$Var(Y_v) = \sum_{i=1}^k a_i^2 Var(y_i) + 2 \sum_{i=1}^k \sum_{j=i+1}^k a_i a_j Cov(y_i, y_j) = S^2,$$

onde k é a quantidade de parâmetros e v o total de combinações lineares (alternativas).

O teste de significância para diferença de médias é dado por:

$$t_c = \frac{\mathcal{E}(Y_{v1}) - \mathcal{E}(Y_{v2})}{\sqrt{\frac{S_1^2 + S_2^2}{N}}}.$$

Como $n_1 = n_2$ o GL (grau de liberdade) é obtido pela fórmula de Aspin-Welch:

$$GL = \frac{(S_1^2 + S_2^2)^2}{\frac{(S_1^2)^2 + (S_2^2)^2}{N+1}} - 2.$$

Para assegurar que o valor global do erro não exceda o nível de significação escolhido, α , deve-se adotar o procedimento de Bonferroni, isto é, cada teste individual deve ser analisado levando-se em conta o nível de significação α/c , onde c é o número de comparações (KNAPP e MILLER III, 1992). O valor calculado é comparado com o valor tabelado, com o nível de significação α/c para o teste t bilateral. Se o valor calculado for maior que o valor tabelado, então existe diferença significativa entre os valores com nível α de significação.

4.2.6. Testes da Independência das Alternativas Irrelevantes – IIA

Um propriedade muito importante relacionado ao modelo Logit Multinomial é a suposição da independência das alternativas irrelevantes.

Seja C um conjunto de alternativas independentes e C_b um subconjunto de C , a probabilidade condicional de uma alternativa $i \in C_b$ em relação a C , devida a independência, é dada por: $P\left(\frac{i}{C_b} \in C\right) = P\left(\frac{i}{C_b}\right) \cdot P\left(\frac{C_b}{C}\right)$, tomando $i, j \in C_b$, a

relação entre as probabilidades condicionais de i e j demonstram a propriedade IIA,

como: $\frac{P\left(\frac{i}{C_b}\right)}{P\left(\frac{j}{C_b}\right)} = \frac{P\left(\frac{i}{C}\right)}{P\left(\frac{j}{C}\right)}$, $i, j \in C_b \subseteq C$ (BEN-AKIVA e LERMAN, 1989).

Mostra que, se as alternativas do conjunto C são independentes, então uma alternativa irrelevante pode ser retirada, restando o conjunto C_b , sem que a relação entre as probabilidades não se alteram. Desse conceito, surge a necessidade da ortogonalidade entre as alternativas, tornando-as independentes e não correlacionadas.

Exemplo: considere-se uma linha de ônibus, servida por dois tipos de transportes coletivos:

Alternativa 1: Micro ônibus Executivo na cor amarela (OE);

Alternativa 2: Ônibus convencional (OC).

Supõe-se uma demanda de 50% para cada tipo de transporte:

OE = 0,5 e OC = 0,5. A razão OE/OC = 0,5/0,5 = 1.

De acordo com a propriedade IIA: “A razão das probabilidades entre 2 alternativas é independente da presença ou ausência de outras alternativas”. Desta forma, se uma nova alternativa de micro ônibus executivo, na cor azul, for incluída, então, de acordo com a propriedade IIA, deveria ter:

OE(Amarelo) = 0,33; OE(Azul) = 0,33; OC = 0,33.

Mas na realidade o que se espera é:

OE(Amarelo) = 0,25; OE(Azul) = 0,25; OC = 0,5, neste caso a propriedade IIA não é satisfeita.

Observação:

Quando a propriedade IIA não é satisfeita, não se recomenda:

- o uso do Modelo Logit Multinomial;
- o uso da divisão do conjunto de alternativas, em conjuntos de escolha;
- o uso de “Rank” explodido.

Quando a propriedade IIA não é satisfeita, recomenda-se o uso do Modelo Logit Hierárquico, com extensão da hierarquia na medida em que a propriedade IIA torne verdadeira.

Para um modelo logit, a razão entre duas alternativas, j e k , no conjunto de

escolha é dada por:
$$\frac{P_{ij}}{P_{ik}} = \frac{e^{(V_{ij})}}{e^{(V_{ik})}} = e^{[X_i'(\beta_j - \beta_k)]}$$

A idéia do teste de partição do conjunto de escolha é simples: se a propriedade IIA é válida, então os parâmetros obtidos através do conjunto de escolha completo (C), são estatisticamente iguais aos parâmetros obtidos com um conjunto restrito (D), ou seja, o conjunto $D \subset C$.

a) Teste de Hausman-McFadden (HM):

O teste estatístico HM (HAUSMAN e McFADDEN, 1984) baseiam-se no fato de ser a propriedade IIA válida. Denota-se: $\delta = (\beta_1, \dots, \beta_L)$ os parâmetros obtidos através da maximização da função de log-verossimilhança do modelo logit; para o conjunto completo C por δ_c , e para o conjunto restrito D por δ_D . Esses parâmetros seriam estatisticamente iguais. O Teste HM é dado por:

$$HM = (\delta_D - \delta_c)' \Omega^{-1} (\delta_D - \delta_c) = q' \Omega^{-1} q$$

onde Ω^{-1} é a inversa generalizada da matriz de covariâncias de q . HUSMAN e McFADDEN (1984) demonstram que Ω^{-1} é equivalente a $(\Omega_D - \Omega_C)^{-1}$, a inversa generalizada da diferença das matrizes de covariâncias de δ_D e δ_c , respectivamente. O teste HM é assintótico com a distribuição Qui-quadrado com graus de liberdade igual ao rank de Ω^{-1} .

b) Teste de McFadden-Train-Tye (MTT):

O teste MTT (McFADDEN et al., 1981) consiste na diferença das estimativas entre o valor da função de verossimilhança de δ_D e δ_c , dado por:

$$MTT = -2[\log L(\delta_c) - \log L(\delta_D)]$$

O teste MTT é assintótico com a distribuição Qui-quadrado com graus de liberdade igual a dimensão de δ_D . Caso o valor calculado seja maior que o valor tabelado, então rejeita-se a hipótese nula, ou seja, a propriedade IIA não é válida.

4.2.7. Elasticidade no Modelo Logit Multinomial

Segundo BEN-AKIVA e LERMAN (1989), a concepção de elasticidade está implícita nas propriedades dos modelos de demanda. A elasticidade para modelo de escolha discreta pode ser distinguida em elasticidade desagregada e elasticidade agregada.

A elasticidade desagregada representa a sensibilidade da propriedade de uma escolha quando tem-se um dos parâmetros alterado. De forma simplificada, a elasticidade seria a alteração provocada na probabilidade da escolha da alternativa i , quando se altera um dos parâmetros k , da variável independente X_{ink} . Neste caso, temos a *elasticidade direta*, $E_{X_{ink}}^{P_n(i)}$, representada pelo produto da derivada parcial de $P_n(i)$ em relação a variável independente X_{ink} , pelo peso do atributo k na probabilidade da alternativa i , $x_{ink}/P_n(i)$.

$$E_{X_{ink}}^{P_n(i)} = \frac{\partial P_n(i)}{\partial x_{ink}} \cdot \frac{x_{ink}}{P_n(i)} = \frac{\partial \ln P_n(i)}{\partial \ln x_{ink}}, \text{ como a derivada parcial}$$

$$\frac{\partial P_n(i)}{\partial x_{ink}} = \frac{\partial}{\partial x_{ink}} \cdot \frac{e^{\beta' x_{in}}}{\sum_i e^{\beta' x_{in}}} = \frac{\sum_i e^{\beta' x_{in}} \cdot e^{\beta' x_{in}} \cdot \beta_k - e^{\beta' x_{in}} \cdot e^{\beta' x_{in}} \cdot \beta_k}{\left(\sum_i e^{\beta' x_{in}} \right)^2} = P_n(i) \cdot \beta_k - (P_n(i))^2 \beta_k$$

$$\text{então } E_{X_{ink}}^{P_n(i)} = [1 - P_n(i)] \cdot P_n(i) \cdot \beta_k \cdot \frac{x_{ink}}{P_n(i)} = [1 - P_n(i)] \cdot x_{ink} \cdot \beta_k$$

Similarmente a elasticidade cruzada, $E_{X_{jnk}}^{P_n(i)}$, da probabilidade da alternativa i quando se altera o valor de um parâmetro de uma outra alternativa j .

$$E_{X_{jnk}}^{P_n(i)} = \frac{\partial \ln P_n(i)}{\partial \ln x_{jnk}} = -P_n(i) \cdot x_{jnk} \cdot \beta_k, \quad \text{para } j \neq i.$$

4.3. TRANSFORMAÇÃO DOS DADOS

4.3.1. Considerações sobre o modelo matemático

Todo delineamento experimental possui um modelo matemático intrínseco e, para poder efetuar a análise de um experimento (obter as estimativas dos parâmetros) é necessário considerar o modelo e formular algumas hipóteses para a validação das estimativas. No caso Logit Multinomial com Probabilidade Condicional, aplicado às técnicas de preferência declarada, o modelo matemático representa a probabilidade associada a uma ordenação das alternativas, realizada pelo entrevistado n , através do processo de escolhas explodidas ($J = 1, 2, \dots, j-1$), no conjunto completo de escolha b , segundo as respectivas utilidade de cada alternativa (BEN-AKIVA e LERMAN, 1989):

$$P_n(b) = \frac{\sum_{i=1}^J e^{\phi V_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\phi V_{inb}}} \times \prod_{j=1}^{J-1} \frac{e^{\phi V_{jn b}}}{\sum_{i \in C_b} e^{\phi V_{inb}}} \times e^{\varepsilon_n} \quad (1)$$

Quando um conjunto de alternativas é submetido ao entrevistado para o processo de escolha explodida, cada escolha é enquadrada em um padrão probabilístico de mensuração dentro da ordenação das escolhas, que nem sempre estaria coerente com a opinião do entrevistado, provocando erro de mensuração das alternativas intra-conjunto.

Em geral, a utilidade de uma alternativa é dada como uma soma de dados observados com um componente aleatório. Assim: $U_{in} = V_{in} + E_{in}$ (BEN-AKIVA e LERMAN, 1989). As hipóteses básicas admitidas para a validade da análise da utilidade U_{in} , onde o erro E_{in} , expresso por $e^{E_{in}}$ em (1), são aditivas, independentes, homocedásticos e com distribuição Gumbel de parâmetro local η e parâmetro escalar $\phi > 0$, figura 4.1.

Segundo BEN-AKIVA e LERMAN (1989) é conveniente atribuir: $\phi = 1$. O parâmetro ϕ não é identificável no modelo, mas isto não obscurece sua compreensão. Ele reflete a suposição de perturbações da homocedasticidade; se esta suposição é violada para a população em questão, a transformação dos dados torna-se necessária para modelar corretamente as utilidades das escolhas.

Um dos casos mais freqüentes de não satisfação das hipóteses básicas é aquele em que não existe homocedasticidade, ou seja a variância não é a mesma nas diferentes alternativas (BANZATTO e KRONKA, 1989). Isto caracteriza a heterocedasticidade ou heterogeneidade dos erros e pode ser de dois tipos.

- *Heterocedasticidade irregular*: ocorre quando a utilidade de uma alternativa apresenta maior variabilidade que as utilidades das outras, sem que haja relação com a média.
- *Heterocedasticidade regular*: ocorre devido à falta de normalidade dos dados experimentais, existindo, freqüentemente, uma certa relação entre a média e a variância das diversas utilidades das alternativas.

A interpretação da probabilidade da utilidade marginal de cada alternativa é dada pela probabilidade da posição em que a alternativa foi indicada pelo entrevistado, n , no conjunto de escolha, considerando as estimativas dos parâmetros. Seja A a utilidade marginal de uma alternativa, calculada pelo produto escalar do vetor das estimativas β^* pelo vetor dos níveis X_A dos respectivos atributos da alternativa correspondente: $A = \beta^* \cdot X_A$. Então, para determinar a probabilidade da utilidade marginal das alternativas no conjunto de escolha definida pelo processo de escolha explodida por um entrevistado n ,

pode ser através da seguinte fórmula recursiva:
$$P_n(A)_K = \prod_{k=1}^K \frac{e^{(A)_k}}{\sum_{i \in J} e^{(A)_i}} \quad (2)$$

K = posição em que a alternativa foi classificada no processo de ordenação;

A_i = é uma das alternativas do conjunto J de alternativas;

J = é o conjunto de alternativas A_i onde foi feito a escolha de A ;

A = é a alternativa classificada na k -ésima ordem; A é alguma A_i .

A seguir calcula-se a média e a variância das probabilidades das utilidades marginais para cada alternativa. As hipóteses básicas para estas estatísticas são de que as médias possuam distribuição normal com parâmetros μ e σ^2 e as variâncias sejam homocedásticas. A figura 4.1 apresenta a função de densidade de probabilidade para a Distribuição Gumbel e as variabilidades das probabilidades das utilidades marginais das alternativas.

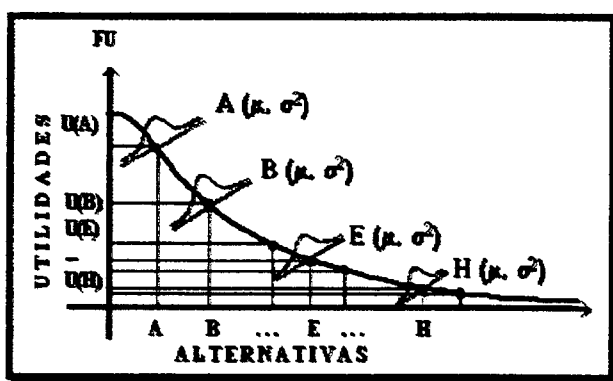


Figura 4.1: Variabilidade das Probabilidades das Utilidades Marginais das Alternativas na Função de densidade de probabilidade da Distribuição Gumbel: $f(x) = e^{(-x - e^{-x})}$ ($\phi=1, \eta=0$)

4.3.2. Teste para homocedasticidade

Um dos testes mais utilizados para verificar a homocedasticidade é o *teste de Hartley* ou *teste da razão máxima*. Considerando um conjunto de g alternativas, cada uma com r repetições, para as quais deseja testar a homocedasticidade. Calcula-se as estimativas das variâncias s^2 , as médias das alternativas e a estatística $H_c = \frac{s_{\max}^2}{s_{\min}^2}$, onde s_{\max}^2 é a maior variância e s_{\min}^2 a menor variância. O valor calculado é comparado com o valor tabelado: se $H_c \geq H_{(g, r-1), \alpha}$, rejeita-se a hipótese de homocedasticidade e conclui-se que existe heterocedasticidade.

Se as alternativas não possuírem a mesma repetição r será representado pela média das repetições. Este teste é eficiente para $g \leq 12$ e todas as alternativas com o mesmo número de repetições. (PEARSON e HARTLEY, 1970). Existem, além deste, outros testes para verificação de homocedasticidade, como o teste de Cochran e o teste de Bartlett, porém, o de Hartley é de aplicação mais simples.

4.3.3. Transformação de dados

Segundo BOX *et al.* (1978) se for constatado que os dados, contrariando as suposições básicas, apresentem uma relação funcional do desvio padrão σ com a média μ , no valor atribuído às alternativas, então é possível buscar uma transformação dos dados $Y = f(y)$ tal que torne a variância constante, corrigindo a heterocedasticidade.

Para obter a potência indicada para a transformação dos dados, calcula-se a equação de regressão entre o logaritmo da média e o logaritmo do desvio padrão de cada alternativa. Como σ e μ são desconhecidos, podemos utilizar suas estimativas. A transformação potência indicada será:

$Y = y^\lambda$. Onde: $\lambda = 1 - \alpha$ e $\text{Ln}(\sigma) = \text{const} + \alpha \text{Ln}(\mu)$.

A tabela 2.3.3, extraído do livro de BOX *et al.* (1978), apresenta um resumo das principais transformações de potência usadas para estabilizar a variância, quando σ está relacionado com μ^α . Para fazer uso da tabela, calcula-se o valor de λ , localiza-se na tabela o valor que mais se aproxima e utiliza-se da transformação indicada.

Tabela 4.3 - Relação de dependência e indicação de transformação

Relação de dependência	α	$\lambda = 1 - \alpha$	Transformação indicada
$\sigma \propto \mu^2$	2	-1	Recíproca
$\sigma \propto \mu^{3/2}$	3/2	-1/2	Raiz quadrada da recíproca
$\sigma \propto \mu$	1	0	Logaritmo
$\sigma \propto \mu^{1/2}$	1/2	1/2	Raiz quadrada
$\sigma \propto \text{const.}$	0	1	Sem transformação

Fonte: BOX *et al.* (1978).

BARTLETT (1947) formulou uma solução teórica para o caso em que existe uma relação entre a média μ e a variância σ^2 e os dados apresentam heterocedasticidade regular, que consiste no seguinte: seja $\sigma^2 = f(\mu) = V(x)$. Generalizando $f(\mu) = f(x)$, busca-se uma transformação $z = g(x)$, tal que a variância $V(z)$ seja constante.

Desenvolvendo $g(x)$ em série de Taylor, em torno da média μ , com aproximação de 1º grau, obtém-se: $z = g(x) = g(\mu) + (x - \mu) g'(\mu)$,
com esperança, $E[z] = E[g(\mu) + (x - \mu) g'(\mu)] = g(\mu)$ e
variância de z , $V(z) = E[z - E[z]]^2 = V(x).[g'(\mu)]^2$.

Como $V(x) = f(\mu) = \sigma^2$, então $V(z) = f(\mu).[g'(\mu)]^2 = k$ (por hipótese, k é uma constante positiva), assim a função $g(\mu)$ pode ser obtida pela equação: $g(\mu) = \int \sqrt{\frac{k}{f(\mu)}} d(\mu)$.

Segundo BEN-AKIVA e LERMAN (1989), o erro no modelo Logit Multinomial tem distribuição Gumbel com parâmetros η e ϕ :

$$F(\varepsilon) = \exp[-e^{-\phi(\varepsilon - \eta)}], \quad \phi > 0.$$

A distribuição tem média $\mu = \eta + \gamma/\phi$, onde γ é a constante de Euler ($\sim 0,577$) e $\eta = 0$

e variância $\sigma^2 = \frac{\pi^2}{6\phi^2} = \frac{\pi^2 \mu^2}{6\gamma^2} = k_1 \mu^2$.

Como $f(\mu) = k_1 \mu^2$, a variância da transformada z , fica: $V[z] = [g'(\mu)]^2 \cdot f(\mu) = k$ (constante) onde $g(\mu) = \int \sqrt{\frac{k}{f(\mu)}} d(\mu)$.

De forma geral, $z = g(x) = \int \sqrt{\frac{k}{f(x)}} dx = \int \sqrt{\frac{k}{k_1 \cdot x^2}} dx = k_2 \cdot \ln|x| + k_3$. Obtendo-se

como resultado, a transformação a ser utilizada, $\ln|x|$, que tornará a variância constante conforme pode ser verificado: $V(z) = k_1 \mu^2 \left(\frac{1}{\mu}\right)^2 = k_1$.

4.3.4. Estimativa de parâmetros no Modelo Logit Multinomial com probabilidade condicional

Seja $P_n(b)$, a probabilidade de uma ordenação das alternativas no conjunto b , efetuada por um entrevistado n , e $\beta'X_j$ o valor da utilidade da alternativa, então:

$$P_n(b) = \frac{\sum_{i=1}^J e^{\beta'X_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta'X_{inb}}} \times \prod_{j=1}^{J-1} \frac{e^{\beta'X_{jnb}}}{\sum_{j \in C_b} e^{\beta'X_{jnb}}} \quad (3)$$

A função de verossimilhança L^* com a transformação logarítmica,

$$L^* = \prod_{n=1}^N \ln|P_n(b)|, \text{ fica:}$$

$$L^* = \prod_{n=1}^N \left[\ln \omega_b + \sum_{j=1}^{J-1} \left(\beta \chi_{jnb} - \ln \sum_{i=1}^J e^{\beta \chi_{inb}} \right) \right] = \prod_{n=1}^N \theta_n, \text{ com } \omega_b = \frac{\sum_{i=1}^J e^{\beta \chi_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta \chi_{inb}}}.$$

Aplicando o logaritmo na função L^* obtém-se a função de log-verossimilhança:

$$L = \sum_{n=1}^N \ln |\theta_n|. \text{ Como os valores de } \theta_n \text{ são negativos, para desenvolver a função de}$$

log-verossimilhança deve-se tomar os valores θ_n em módulo, o que provoca a inversão do sentido de sua concavidade e passa a ter um mínimo em lugar do máximo. Assim, para utilizar o método de Newton-Raphson, com o objetivo de maximizar a função e estimar os parâmetros β , a função L deve ser multiplicada por (-1) :

$$L = - \sum_{n=1}^N \ln |\theta_n|.$$

A derivada primeira em relação a β_k é $[\nabla L]_k = - \sum_{n=1}^N \frac{\partial \theta_{nk}}{\theta_n}$, e a derivada segunda em

relação a β_l fica $[\nabla^2 L]_{kl} = - \sum_{n=1}^N \frac{(\theta_n)(\partial^2 \theta_{nkl}) - (\partial \theta_{nk})(\partial \theta_{nl})}{\theta_n^2}$. Dessa forma

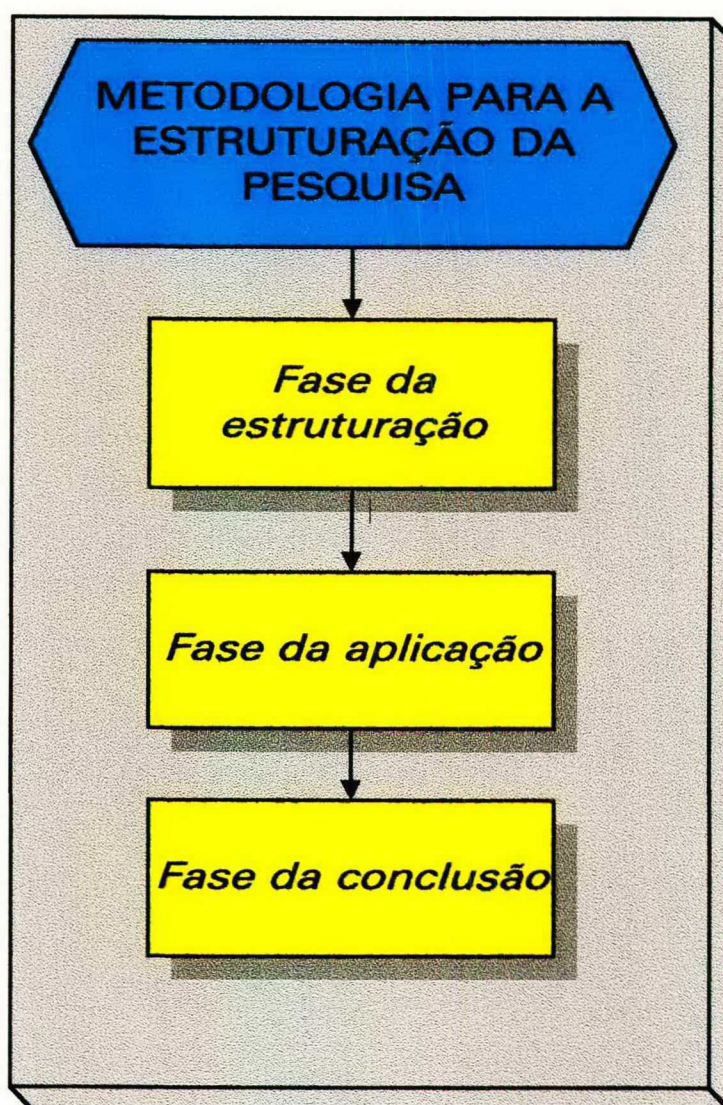
será possível estimar os parâmetros, β^* , para os dados transformados.

4.4. IMPLEMENTAÇÃO COMPUTACIONAL – SOFTWARE LMPC

Todos os procedimentos e testes descritos nesta seção estão implementados no software denominado LMPC – Logit Multinomial com Probabilidade Condicional, desenvolvido em ambiente Delphi (linguagem pascal), pelo autor da tese.

PARTE I – ASPECTOS TEÓRICOS

Capítulo 5 - METODOLOGIA PARA A ESTRUTURAÇÃO DA PESQUISA



5. METODOLOGIA PARA A ESTRUTURAÇÃO DA PESQUISA

Para a montagem de um delineamento experimental com a abordagem da pesquisa de preferência declarada, pode-se organizar os procedimentos em 3 fases:

- Estruturação;
- Aplicação;
- Análise e Interpretação.

5.1. FASE DA ESTRUTURAÇÃO

5.1.1. Identificar o objetivo geral da pesquisa de preferência declarada

A primeira etapa na montagem de uma pesquisa de campo é a definição do objetivo geral a ser atingido, ou seja, equacionar bem o problema, levantar os objetivos operacionais e a finalidade da pesquisa de campo. O pesquisador deve traçar planos e metas para atingir os objetivos e seguir uma metodologia adequada para o tipo de pesquisa de campo a ser adotada.

5.1.2. Identificar os objetivos operacionais e as áreas de interesse

Identificar os objetivos operacionais de tal forma a obter os meios para se chegar ao objetivo geral. Para identificar os objetivos operacionais, bem como os atributos a serem inseridos na pesquisa, podem ser utilizados dois métodos:

- a. inserindo os atributos identificados como relevantes, em uma pesquisa de opinião realizada com tal finalidade, ou
- b. montando a pesquisa com base nos objetivos e atributos definidos pela pessoa ou entidade que a encomendou.

No primeiro caso, realiza-se uma pesquisa através de questionários abertos em que cada entrevistado cita os atributos mais importantes e, através de métodos descritivos, identificam-se os atributos relevantes para a população a ser amostrada, por exemplo os estudos realizados por BASTOS (1994) e MAY (1996).

No segundo caso, o processo toma uma forma mais iterativa entre as pessoas envolvidas: Administrador, Analista e Entrevistado.

- ADMINISTRADOR – quem encomenda a pesquisa e define os objetivos e as áreas de interesse para a situação em estudo;
- ANALISTA – quem organiza, orienta e analisa a pesquisa;
- ENTREVISTADO – público alvo da pesquisa.

Uma das primeiras atividades a ser desenvolvida pelo Analista é a de identificar, junto ao Administrador, o objetivo geral, os objetivos operacionais e as áreas de interesse. Dentro de cada área de interesse, identificar os atributos relevantes objetos de interesse do Administrador.

Os **atributos**, aqui definidos, são as dimensões de uma área de interesse, em que se pretende estabelecer uma Função Utilidade local (FU_l) da escala de preferência do público alvo amostrado.

O Analista deve, através de seções iniciais de “brainstorming”, levantar, junto ao Administrador, uma lista de elementos iniciais que se candidatarão a atributos de avaliação. Essa lista de elementos, a princípio, aparentar-se-á de forma caótica, desconexa e mal definida; o Analista deve, então, clarificá-las, torná-las operacionais, encontrar suas interconexões e incompatibilidades, enfim, estruturá-la (BANA E COSTA, 1992).

O processo de estruturação consiste em dispor os elementos iniciais em um diagrama de correlação hierárquica, formando uma árvore ramificada, cujo tronco central representa o objetivo geral e cada ramo principal se constitui numa área de interesse; cada área de interesse é constituída por vários atributos e cada atributo por seus níveis, conforme a figura 2.4.1. Esse processo pode ser aplicado a ambos os métodos descritos acima.

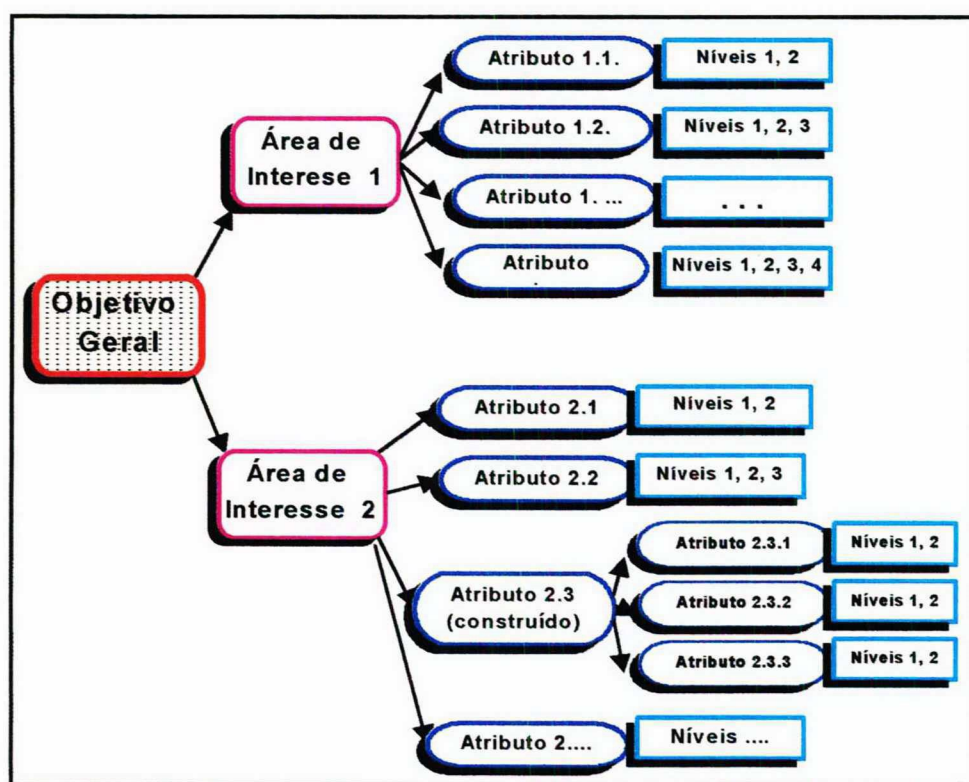


Figura 5.1. Árvore da organização do experimento

Cada área de interesse poderá constituir-se num experimento de pesquisa de preferência declarada. Para tanto, há necessidade do pressuposto de que os atributos sejam ortogonais. A ortogonalidade é entendida no sentido da mínima correlação entre os atributos.

Por exemplo:

Atributo A: rapidez na passagem pelos caixas de um supermercado;

Atributo B: número de caixas no supermercado;

Atributo C: Iluminação no interior do supermercado.

Podemos perceber que entre os atributos A e B existe uma correlação muito forte, ou seja, a rapidez na passagem dos clientes pelos caixas depende, entre outras, do número de caixas disponíveis; logo, não são ortogonais. Entre os atributos A e C ou B e C a correlação é mínima, ou seja a “rapidez na passagem” tem pouca dependência com a “iluminação interna” e o “número de caixas” não tem dependência com a “iluminação interna”; logo, podemos considerar como atributos ortogonais.

5.1.3. Identificar e definir os atributos relevantes em cada área de interesse.

Para melhor uniformizar o entendimento do experimento, deve-se definir bem as **áreas de interesse**, de tal sorte que elas explicitem os objetivos operacionais e propiciem a obtenção do objetivo geral.

Para cada **atributo**, deve-se conceituar o **pólo presente** (isto é, definir um rótulo para a situação atual ou ideal) e o **pólo contrário** (isto é, definir um rótulo para a situação contrária). Por exemplo:

Atributo C: iluminação no interior do supermercado (iluminação adequada “que permita ler perfeitamente as embalagens dos produtos” ... iluminação fraca “a leitura é feita com dificuldade”).

Os dois rótulos são separados por “...” (lido como “ao invés de”). O texto de cada conceito não deve ser muito longo (aproximadamente 12 palavras), o mais abreviado possível, buscando-se manter as palavras e frases usuais para o caso. O sentido dos atributos é obtido através do contraste entre os dois pólos (ACKERMAN et al., 1995).

5.1.4. Construir e descrever cada nível dos atributos;

O conjunto de situações ou estados definidos entre os pólos dos atributos se constitui nos **níveis do atributo**. Os níveis do atributo são inseridos no escopo dos seus pólos.

BANA E COSTA (1992) propõe três dimensões para os níveis dos atributos:

- quantitativos ou qualitativos;
- discretos ou contínuos;
- diretos, indiretos ou construídos.

a) Quantitativos ou qualitativos

- O nível é quantitativo quando pode ser descrito por números, ou seja, os valores quantitativos descrevem adequadamente os seus diversos níveis;

- O nível é qualitativo quando os números não descrevem adequadamente os seus níveis e há necessidade da introdução de expressões semânticas para descrevê-los. Exemplo: “acima das minhas expectativas, na expectativa, abaixo das minhas expectativas”.

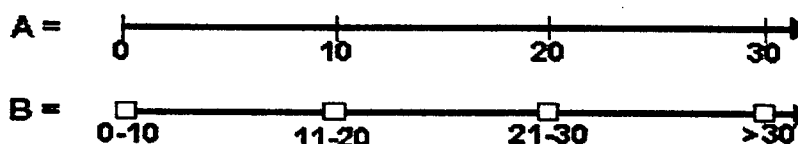
b) Discretos ou Contínuos

- O nível do atributo é considerado discreto quando os níveis possíveis são enumeráveis;
- O nível do atributo é considerado contínuo quando podemos relacionar seus níveis a uma função matemática contínua.

Quando o atributo possui níveis contínuos, pode-se criar níveis discretos da seguinte forma:

Onde: - Níveis de A: 0, 10, 20 e 30 e

- Níveis de B (intervalos): (0, 10), (11, 20), (21, 30), (31,...).



Para a incorporação de atributo contínuo, em um delineamento com arranjos fatoriais, deve-se discretizar o escopo dos valores possíveis, em níveis (3 ou 4 níveis), e atribuir:

- ao valor base o Nível Base (NB),
- ao valor superior o Nível Superior (NS), e
- no intervalo, dividir em 2 ou 3 partes ou Níveis Intermediários (NI).

A tabela 5.1 apresenta uma forma de codificar os níveis dos atributos discretos. Cabe ressaltar que, para atributos com 3 ou 4 níveis, devem ser utilizados dois parâmetros para definir os valores entre os níveis, ou seja, há a necessidade de dois dígitos na codificação binária: 00, 01, 10 e 11. Essas codificações corresponderiam, respectivamente, aos níveis: 0, 1, 2 e 3.

No caso de o atributo possuir 3 níveis discretos, estes deverão ser designados a três dos 4 níveis de códigos binários; o quarto nível será designado simplesmente através da repetição de um dos três níveis anteriores (aquele que achar melhor) e a ortogonalidade é mantida (ROSS, 1991).

Tabela 5.1 – Codificação para atributos com níveis discretos

Atributos com 2 Níveis		Atributos com 3 Níveis		Atributos com 4 Níveis	
Níveis	Códigos	Níveis	Códigos	Níveis	Códigos
0 – NB	0	0 – NB	0 0	0 – NB	0 0
1 – NS	1	1 – NI	0 1	1 – NI ₁	0 1
		2 – NS	1 0	2 – NI ₂	1 0
				3 – NS	1 1

A codificação dos níveis de um atributo contínuo requer cuidados especiais. Para sua correta utilização, primeiro deve-se definir os valores para os níveis NB e NS e em seguida, determinar a proporcionalidade entre os níveis intermediários. Os valores ou códigos a serem incluídos no delineamento devem ser iguais ou proporcionais aos valores dos níveis reais. Em alguns casos, há a necessidade da transformação dos níveis. A tabela 5.2 apresenta alguns exemplos de codificação dos níveis contínuos dos atributos.

Tabela 5.2 – Exemplos de Codificação para atributos com níveis contínuos

Atributos com 2 Níveis		Atributos com 3 Níveis				Atributos com 4 Níveis			
Níveis	Códigos	Níveis	Códigos			Níveis	Códigos		
0 – NB	0	0 – NB	0	-1	0	0 – NB	0	-3	0
1 – NS	1	1 – NI	1	0	0,5	1 – NI ₁	1	-1	0,33
		2 – NS	2	1	1	2 – NI ₂	2	1	0,67
						3 – NS	3	3	1

Caso haja interesse no teste t de comparação entre os valores de utilidade das alternativas (teste descrito no item 4.2.5), há a necessidade de codificar os níveis NB e NS em níveis binários (0 1), respectivamente, e estabelecer a proporcionalidade para os níveis intermediários. Tal procedimento deve ser uniforme em todos os atributos, sejam contínuos ou discretos.

c) Diretos, Indiretos ou Construídos

- Um nível é *direto* quando está diretamente associado ao atributo. Exemplo: Atributo: distância => Níveis: Km;
- Um nível é *indireto* quando devem-se lançar mão de valores indiretos para definir os níveis. Exemplo: Atributo: distância => Níveis: tempo (quando para medir a distância, utiliza-se uma medida de tempo).
- Um nível é *construído*, quando utilizam-se outros subatributos e níveis correlacionados para construí-lo. Exemplo: Atributo 2.3 da figura 5.1; Atributo: Qualidade do percurso => Níveis: composição dos subatributos (tempo de percurso, espaço em Km, Tipo de percurso, etc.), nesse caso, é conveniente que os subatributos sejam de dois subníveis.

Uma exigência básica que deve ser imposta ao nível do atributo é que ele operacionalize adequadamente o atributo e não seja ambíguo. O nível que tem significado claro e preciso é considerado não ambíguo. KEENEY (1992) apresenta três propriedades desejáveis:

- **Mensurabilidade** – quando define perfeitamente o atributo, níveis do tipo *bom*, *fraco*, etc. prejudicam a mensurabilidade;
- **Operacionalidade** – quando descreve bem as situações e serve como base sólida para a realização de julgamentos de valor;
- **Compreensibilidade** – quando os níveis dos atributos não permitem ambigüidade nem permitem perda de informações (termos ambíguos: eficácia, eficiência, competitividade, melhoria, capacidade, ou então, níveis do tipo: nenhum, mínimo, baixo, moderado, alto, que expressam julgamento subjetivo de valor).

Se puder escolher os níveis dos atributos recomenda-se os **diretos, contínuos e quantitativos**. Níveis de atributos construídos avaliam bem as dimensões relacionadas ao atributo mas são suscetíveis a problemas de compreensibilidade e operacionalidade. Níveis de atributos indiretos tornam a tarefa mais simples mas aumentam o risco de redundância e erros de significação.

5.1.5. Identificação das combinações possíveis com os níveis dos atributos

Cada combinação possível de níveis e atributos que possam ser implantados na realidade, recebe o nome de alternativa. Antes de construir o esquema fatorial que resultará nos grupos de alternativas a serem submetidas aos entrevistados para o processo de escolha, pode ser necessário excluir as alternativas irrelevantes ao caso (BEN-AKIVA e LERMAN, 1989), ou seja, excluir as alternativas em condições fora da realidade e as alternativas, já definidas "a priori", como dominantes ou totalmente dominadas.

No entanto, é conveniente que as alternativas excluídas por serem consideradas dominadas ou dominantes façam parte do delineamento no processo de estimativa dos parâmetros. No caso, dominada, será considerada como última escolha, e dominante, será considerada como primeira escolha. Assim, os parâmetros apresentam um comportamento uniforme e coerente, caso contrário, a matriz hessiana poderá ser inversível.

a) Preferência do Consumidor

Segundo VINCKE (1992), quando o entrevistado for fazer a comparação entre duas alternativas (a e b), de início, ele agirá de acordo com uma das três situações:

- (1) preferência por uma das alternativas;
- (2) indiferença entre as alternativas;
- (3) incomparabilidade entre as alternativas.

Essa relação de preferência deve ser entendida como uma noção operacional. Se o consumidor prefere uma alternativa à outra, isso significa que sempre que tiver oportunidade ele escolherá a primeira em vez da segunda. Portanto, a idéia de preferência está baseada no **comportamento** do consumidor. Para poder dizer que uma alternativa é preferida à outra, nós observamos como o consumidor se comporta em situações de escolha envolvendo as duas alternativas. Se ele escolhe a alternativa a quando b podia também ser escolhida, então é natural que esse consumidor prefere a à b . Desse princípio surge a necessidade de, sempre que possível, confrontar todas as alternativas duas a duas.

b) Hipóteses sobre as Preferências

São feitas, usualmente, algumas suposições sobre a “consistência” das preferências do consumidor. Algumas das suposições são tão fundamentais que serão chamadas de “axiomas” da teoria do consumidor.

Eis três axiomas sobre as preferências do consumidor.

Completa: duas alternativas quaisquer podem ser comparadas.

Reflexiva: qualquer alternativa é ao menos tão boa quanto ela mesma.

Transitiva: se a alternativa *a* é ao menos tão boa quanto *b* e se *b* é ao menos tão boa quanto *c*, então o consumidor acha que *a* é ao menos tão boa quanto *c*.

O primeiro axioma, de que as preferências são completas, é dificilmente questionável, ao menos para o tipo de escolhas em pesquisas de marketing. Dizer que duas alternativas podem ser comparadas é dizer simplesmente que o consumidor é capaz de fazer uma escolha entre as alternativas.

O segundo axioma, reflexividade, é trivial. Qualquer alternativa é certamente ao menos tão boa quanto uma alternativa idêntica.

O terceiro axioma, *transitividade*, é mais problemático. Não é claro que a transitividade seja necessariamente uma propriedade que as preferências devam satisfazer. A hipótese das preferências serem transitivas não parece obrigatória em termos de lógica pura, como de fato ela não é. Transitividade é uma hipótese sobre o comportamento de escolha das pessoas, não um enunciado de lógica pura. Se essa hipótese é ou não um fato básico de lógica, não é o ponto interessante: o que interessa é se ela representa ou não uma descrição adequada do modo como as pessoas se comportam.

5.1.6. Organizar o esquema fatorial adequado para cada área de interesse

O esquema fatorial a ser adotado depende de alguns fatores:

- do pressuposto da ortogonalidade entre os atributos, o que possibilitará a montagem de delineamento de arranjos ortogonais;
- do número de atributos e do número de níveis em cada atributo.

Este trabalho apresenta algumas possibilidades de Planos Experimentais. O anexo 1 apresenta as indicações de delineamentos experimentais, baseados nos arranjos ortogonais de Taguchi (TAGUCHI, 1988) e (ROSS, 1991) e nos arranjos fatoriais, através das técnicas de confundimento ou fatorial fracionários (MANN, 1949), (KEMPTHORNE, 1967), (COCHRAN e COX, 1978), (BOX, HUNTER e HUNTER, 1978) e (MONTGOMERY, 1984).

A divisão do conjunto de alternativas em blocos, ou seja, conjunto de escolha com probabilidade condicional, baseado nas técnicas de blocos incompletos, deverá ser feita de acordo com a tabela A4.1 do anexo 4.

5.2. FASE DA APLICAÇÃO

- 1) Sempre que possível, organizar e aplicar uma pesquisa piloto para cada área de interesse e realizar os testes de validação da amostra e ajustes;
- 2) Organizar e aplicar a pesquisa de preferência declarada de acordo com cada área de interesse;
- 3) Organizar uma pesquisa “especial” de preferência declarada, se for o caso, para identificar o peso relativo (trade-off) entre as áreas de interesses.

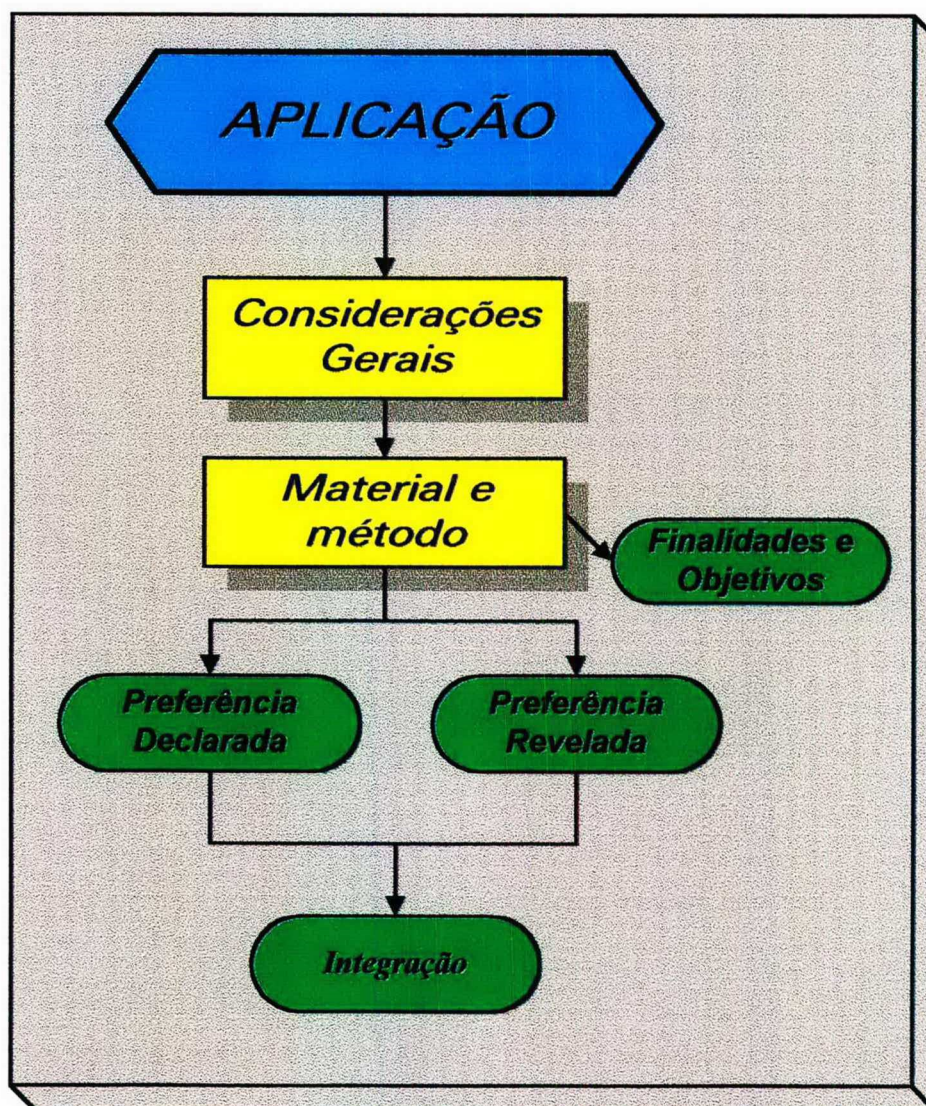
5.3. FASE DA CONCLUSÃO

- 1) Calcular as estimativas dos parâmetros;
- 2) Realizar os testes de validação interna;
- 3) Analisar, interpretar e concluir a pesquisa.

PARTE II – APLICAÇÃO

Capítulo 6 – PESQUISA DE CAMPO:

A utilidade dos supermercados na visão dos clientes



6. PESQUISA DE CAMPO:

A utilidade dos supermercados na visão dos clientes

6.1. CONSIDERAÇÕES GERAIS

6.1.1. O relacionamento com os Clientes

Segundo LEVITT (1990), “relationship” é a palavra que vem substituir, nos anos 90, termos clássicos como posicionamento de produtos, segmentação de mercado e outros, ou seja, o grande instrumento da diferenciação estratégica será o relacionamento.

O relacionamento, tal como um produto, tem aspectos tangíveis, visíveis e claramente definidos como direitos e deveres de fornecedores e clientes. Mas há também os aspectos intangíveis, não identificados com facilidade, que necessitam, muitas vezes, da ajuda de pesquisa para a sua identificação (COBRA e RANGEL, 1993). Dessa maneira, um bom relacionamento só se constrói com base em uma gama de bons serviços ao cliente. COBRA e RANGEL (1993) define serviço como uma forma de proporcionar tantas satisfações quantas possíveis pela posse de um bem ou do serviço adquirido. Um bom serviço é aquele que vai ao encontro das expectativas do cliente ou consumidor. O serviço é uma forma de ampliar um produto vendido.

A qualidade de um serviço, como é percebida pelos consumidores, tem duas dimensões: uma técnica e outra funcional (figura 6.1).

- A dimensão técnica diz respeito ao processo de produção do serviço. Esse processo envolve a utilização de máquinas e equipamentos para a racionalização dos serviços de atendimento.
- A dimensão funcional da qualidade de um atendimento diz respeito à maneira como o consumidor percebe e valoriza a qualidade do serviço.

É fácil observar-se que a avaliação da dimensão funcional está vinculada aos valores subjetivos dos clientes e não pode ser avaliada com a mesma objetividade com que é possível avaliar a qualidade na dimensão técnica. Portanto, a qualidade é

freqüentemente associada a uma das chaves de sucesso empresarial. A vantagem competitiva de uma empresa depende da qualidade de seus próprios produtos e dos serviços prestados (COBRA e RANGEL, 1993).

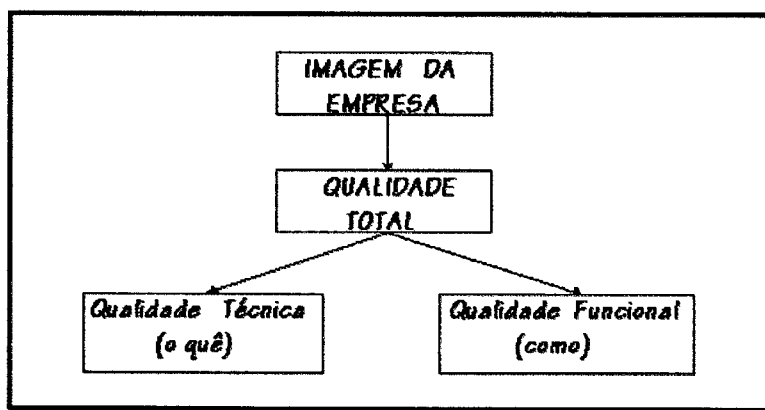


Figura 6.1 – Imagem da Empresa

Fonte: Adaptado de Christian Grondroos, *Service Management and Marketing*. Lexington Books, Lexington, Massachusetts, 1990, apud COBRA e RANGEL (1993).

6.1.2. Características da Cidade de Guarapuava

A cidade de Guarapuava situa-se na região centro sul do Estado do Paraná, no Terceiro Planalto. Com seus mais de 180 anos, foi formada como ponto de passagem de tropeiros, recebendo forte imigração da população de origem germânica no início deste século. Irradia sua influência sobre 11 municípios vizinhos. É um centro industrial vinculado ao reflorestamento, com especialização nos setores de madeira, papel e papelão, tem também importantes indústrias ligadas à de cervejaria. Possui um mercado intenso na área de Cooperativas. Principais rendas: agricultura, pecuária, comércio e indústria de produtos agropecuários. Ocupa uma posição estratégica no estado, reforçada pelo entroncamento rodoviário criado pelas estradas BR-373, BR-277 e pelas rodovias estaduais PR-170 e PR-460. Possui uma população de aproximadamente 170 mil habitantes, dos quais 145 mil residem na zona urbana e 25 mil na zona rural.

6.1.3. Características do sistema de comércio nos supermercados da cidade

Os supermercados que operam na região de Guarapuava possuem características peculiares regionais e, para melhor interpretá-las, foram classificados em dois tipos

administrativos distintos, polarizados conforme descritos na tabela 6.1 – Quadro comparativo entre os tipos de comércios em supermercados:

Tabela 6.1 – Quadro comparativo entre os tipos de comércios

Tipo de comércio “A”	Tipo de comércio “B”
Os produtos são bem organizados nas prateleiras (sem as caixas de papelão);	Os produtos são razoavelmente organizados nas prateleiras (a maioria dos produtos ficam expostos em suas embalagem de papelão);
Há boa diversificação dos produtos à disposição do consumidor;	Só os produtos mais rotativos se encontram à disposição do consumidor;
O número de funcionários, repositores, caixas e empacotadores é suficiente para dar um atendimento de qualidade ao cliente;	O número de funcionários, caixas e repositores é reduzido e não possui serviço de empacotamento de compras;
Aceitam cartões de créditos, cheque pré-datado e outras formas de pagamento a prazo;	As vendas são feitas na modalidade sempre "a vista".
Praticam o serviço de entrega em Domicílio gratuito ou a custo simbólico;	O serviço de entrega em domicílio tem seus custos mais realistas;
Os preços dos produtos em geral são os praticados no mercado, onde estão incluídos os custos dos serviços ao cliente.	Os preços dos produtos em geral são reduzidos, abaixo do mercado. A redução do custo com os cortes dos serviços ao cliente é repassada na forma de redução no preço dos produtos.

A tabela 6.2 apresenta um quadro com a classificação de 11 lojas de supermercados que operam na cidade de Guarapuava e suas respectivas classificações, de acordo com o tipo, que foram inicialmente selecionadas para estudos. Os nomes dos supermercados foram substituídos por códigos, tendo em vista que esta pesquisa científica tem por finalidade desenvolver e aplicar uma metodologia de pesquisa de preferência declarada de forma acadêmica; no entanto, os dados e parâmetros estimados representam a situação real em relação à população amostrada.

Exemplos de identificação dos códigos dos supermercados:

- SM-A1 → supermercado do tipo A de número 1;
- SM-A2 → supermercado do tipo A de número 2;
-
- SM-B1 → supermercado do tipo B de número 1;
- SM-B2 → supermercado do tipo B de número 2;

Tabela 6.2 – Classificação dos Supermercados

Super-mercados	Características dos supermercados	Lojas	Tipo A	Tipo B
SM-A1	Supermercado de grande porte, bem aparelhado, construção recente, com cerca de 2 anos de funcionamento. Está localizado no centro da cidade.	1	X	
SM-A2	Supermercado de grande porte, bem aparelhado, construção recente, com cerca de 5 anos de funcionamento. Está localizado próximo do centro da cidade	1	X	
SM-A3	Supermercado de médio porte, com cerca de 10 anos de funcionamento. Tem 2 lojas localizadas na região central da cidade, próximo o terminal rodoviário urbano.	2	X	
SM-A4	Supermercado de médio porte, com mais de 15 anos de funcionamento. Está localizado na região central da cidade.	1	X	
SM-A5	Supermercado de médio porte, com mais de 10 anos de funcionamento. Está localizado num dos bairros da cidade.	1	X	
SM-A6	Supermercado de médio porte, com mais de 10 anos de funcionamento. Está localizado num dos bairro da cidade.	1	X	
SM-A7	Supermercado de médio porte, com mais de 10 anos de funcionamento. Está localizado num dos bairro da cidade.	1	X	
SM-B1	Supermercado de médio porte, com mais de 10 anos de funcionamento. Tem duas lojas localizadas diagonalmente em relação ao centro.	2		X
SM-B2	Supermercado de médio porte, com mais de 10 anos de funcionamento. Tem uma loja localizada próximo do centro e outra num dos bairros.	2		X

6.1.4. Avaliação dos atributos relativos aos Supermercados

Com a finalidade de testar a validação externa dos resultados obtidos através da pesquisa de preferência declarada (Função Utilidade), foram avaliados os atributos, de acordo com a codificação dos respectivos níveis, referentes aos delineamentos de *qualidade e conveniências*, em 4 lojas de supermercados do tipo A e 4 lojas do tipo B, que operam em Guarapuava, cujos resultados constam da tabela 6.3. Cabe ressaltar que essa avaliação é estritamente didática e subjetiva, mas revestida com o compromisso da realidade. Para os níveis dos atributos referentes ao delineamento *preço*, adotou-se uma pesquisa de preço, ou seja, foram eleitos 5 produtos de cada setor: carnes, mercearia, frutas e verduras, frios e laticínios; em seguida, calculou-se a variação do custo entre os supermercados indicados e obtiveram-se os respectivos descontos em relação ao preço mais alto. Tanto a avaliação dos atributos, em cada supermercado, quanto a pesquisa de preço foram executadas, pelo pesquisador, antes da realização das entrevistas, isto é, na

fase de preparação e organização da pesquisa e serviram de subsídios para a definição dos níveis dos atributos.

As avaliações dos atributos referentes aos principais supermercados que operam na Cidade, de acordo com o tipo e que atendem aos propósitos da pesquisa, listadas na tabela 6.3, serão utilizadas para a obtenção das respectivas utilidades de cada supermercado, do ponto de vista dos clientes, de forma geral e segmentados por renda e permitirão estabelecer uma integração com os dados de preferência revelada, e assim, obter o *Coeficiente de Atratividade do Supermercado (CAS)*, ou seja, o coeficiente que representa, para cada supermercado, o conjunto dos atributos não incluídos na pesquisa de preferência declarada.

Tabela 6.3 – Avaliação Subjetiva dos Atributos

Atributos	Tipo A (4 lojas)			Tipo B (4 lojas)	
	SM-A1 (1)	SM-A2 (1)	SM-A3 (2)	SM-B1 (2)	SM-B2 (2)
<i>Carnes</i>					
Qualidade	A (1,0)	A (1,0)	A (1,0)	C (0,8)	E (0,2)
Preço Médio	0	-1%	0	-5%	0
<i>Bens de Mercearia</i>					
Qualidade	A (1,0)	A (1,0)	B (0,8)	C (0,6)	D (0,4)
Preço Médio	0	-11%	-11%	-17%	-15%
<i>Frutas e verduras</i>					
Qualidade	A (1,0)	A (1,0)	A (1,0)	D (0,8)	D (0,6)
Preço Médio	-10%	-10%	0	-30%	-15%
<i>Frios e Laticínios</i>					
Qualidade	A (1,0)	A (1,0)	B (0,8)	C (0,6)	D (0,4)
Preço Médio	-3%	-5%	0	-21%	-17%
<i>Conveniência</i>					
Estacionamento	1	0,7	1	0,3	0,3
Lojas Conveniência	1	0	0	0	0
Organização	1	1	1	0,6	0,3
Sist. Crédito	1	1	1	0	0
Níveis Serviços	1	1	1	0	0
Diversificação	1	1	0,8	0,5	0,6
Compensação	0	0	0	1	1

6.2. MATERIAL E MÉTODO

6.2.1. Objetivos e finalidades

A pesquisa de campo abrange os serviços oferecidos pelas Redes de Supermercados na Cidade de Guarapuava, PR, onde se terá a oportunidade de aplicar a estrutura dos arranjos ortogonais, bem como o modelo de estimação de parâmetros através da “logit multinomial” com probabilidade condicional, desenvolvidos neste trabalho.

A pesquisa atenderá aos seguintes objetivos:

- a) Determinar as funções matemáticas que expressem a utilidade dos supermercados do ponto de vista da população amostrada, em geral e por classes de rendas;
- b) Identificar a relação de importância entre os atributos que influenciam os clientes na escolha de um supermercado para fazer as compras mensais;
- c) Identificar a porção da população que determinado tipo de supermercado atende;
- d) Validar a aplicação dos testes, técnicas e metodologias desenvolvidas nesta tese, em pesquisa na área de marketing;
- e) Estabelecer um estudo comparativo entre as técnicas de preferência declarada e preferência revelada em pesquisa na área de marketing.

6.2.2. Organização da pesquisa de campo

A organização da pesquisa de campo obedecerá ao organograma da figura 6.2 de forma a identificar: a relação entre os atributos que os clientes levam em consideração na escolha de um supermercado para fazer as compras, através dos dados de preferência declarada; a utilidade dos supermercados em função das preferências realizadas, com dados de preferência revelada; integrando as duas metodologias na busca do coeficiente de atratividade dos supermercados.

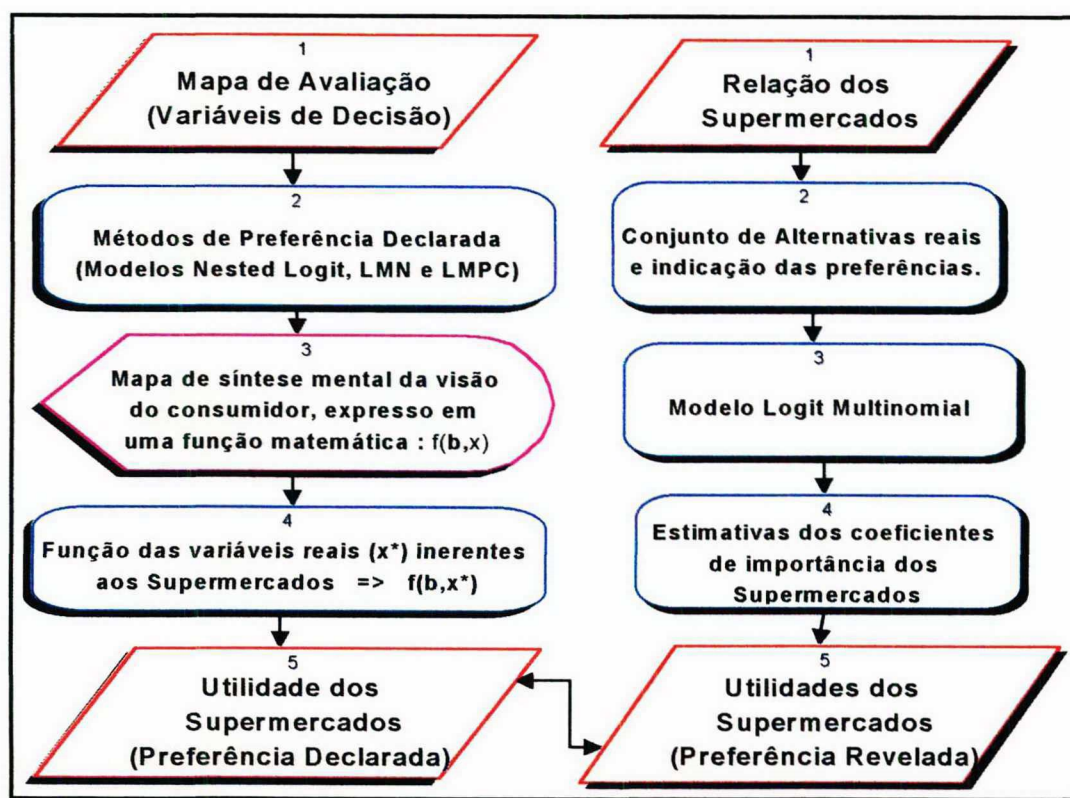


Figura 6.2 – Organograma da Aplicação (Pesquisa de Campo)

6.2.3. Forma de complexidade dos delineamentos experimentais na pesquisa de preferência declarada (PD).

a) Estrutura da pesquisa de preferência declarada.

Na aplicação, os atributos de decisão estarão organizados em áreas de interesses, conforme a figura 6.3, formando os mapas de síntese mental a respeito das variáveis latentes que os consumidores geralmente usam para tomar decisões sobre o local de compras mensais. A abordagem decomposicional terá como objetivo modelar em termos de função utilidade, as variáveis latentes: conceito de preço, qualidade/organização e conveniência, identificadas pelos rótulos de PREÇO, QUALIDADE e CONVENIÊNCIA, respectivamente, a partir da avaliação subjetiva das alternativas, dada uma regra de composição, e assim, estabelecer um julgamento global do ponto de vista dos consumidores. Por exemplo, LOUVIERE e MEYER (1981) apud. (LOUVIERE e GAETH, 1987) demonstram que as variáveis latentes: conceito de preço, qualidade/organização e julgamento de conveniência dos supermercados são altamente preditivos para a escolha do supermercado e que indicam os interesses

definidos pelos questionamentos: Como os preços são bons?; Como a qualidade é boa?, e Como este lugar é conveniente para fazer compras?.

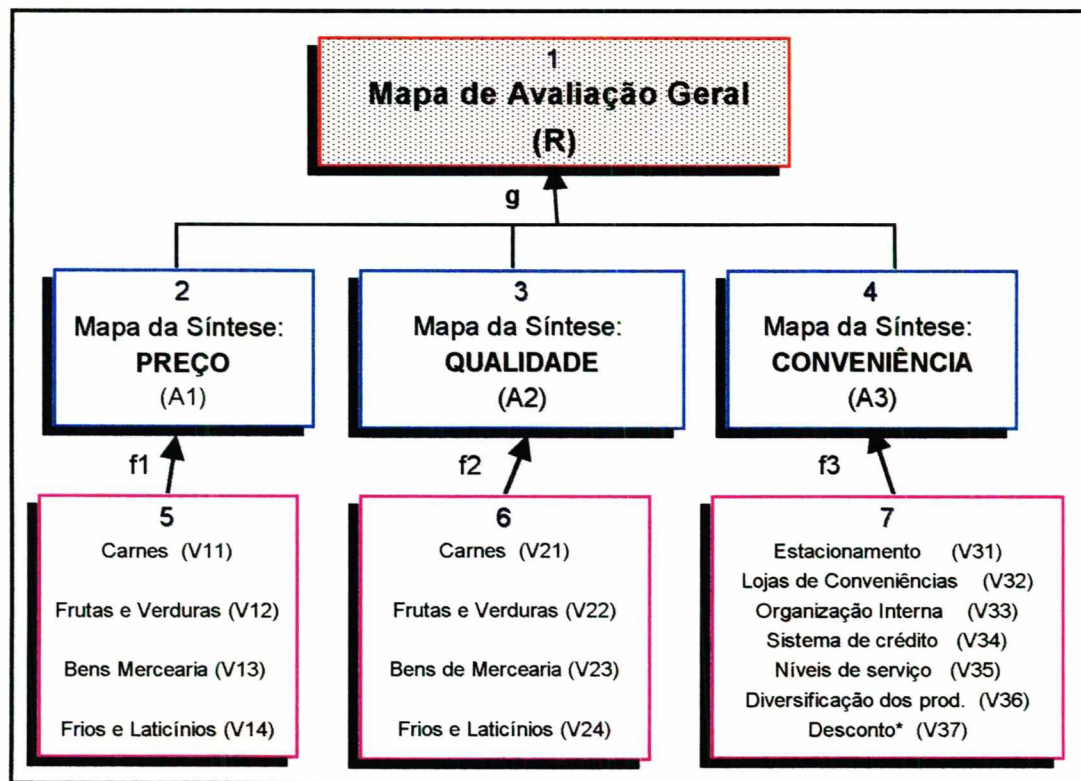


Figura 6.3 – Mapa de Avaliação Geral

Denota-se:

A_i , $i = 1, 2$ e 3 definem os conjuntos dos atributos para cada área de interesse;

V_{ij} , definem os conjuntos dos níveis dos atributos em cada área;

f_i , definem as funções utilidades para cada área de interesse de A_i ;

g é uma função que define o mapa de avaliação geral, R .

$A_1 = \{V_{11}, \dots, V_{1n_1}\}$, $A_2 = \{V_{21}, \dots, V_{2n_2}\}$; $A_3 = \{V_{31}, \dots, V_{3n_k}\}$, são as variáveis que definem cada decisão nos subconjuntos.

$R(A_1), \dots, R(A_3)$ o espaço de resposta associado com cada subconjunto de atributo A_i , respectivamente.

Seja f_i ($i = 1, 2, 3$) funções das variáveis do i -ésimo subconjunto $\{A_i\}$ para o único subconjunto dimensional do espaço, $R(A_i)$.

Então $f_i : \{V_{i1}, \dots, V_{ink}\} \rightarrow R(A_i)$, $i = 1, \dots, 3$.

Agora seja g uma superfície de respostas formada pela composição proporcionais do conjunto de variáveis dos subconjuntos $\{R(A_1), R(A_2), R(A_3)\}$ no espaço de resposta final R .

$g : \{\beta_1.R(A_1); \beta_2.R(A_2); \beta_3.R(A_3)\} \rightarrow R$, com β_i os coeficientes de proporcionalidade entre as áreas de interesses.

Então a superfície de resposta da composição de g com f , pode ser descrita como:

$$\begin{aligned} g \circ f_i [(V_{11}, \dots, V_{1n_1}); (V_{21}, \dots, V_{2n_2}); (V_{31}, \dots, V_{3n_k})] &= \\ &= g [f_1(V_{11}, \dots, V_{1n_1}); f_2(V_{21}, \dots, V_{2n_2}); f_3(V_{31}, \dots, V_{3n_k})] = \\ &= \beta_1.R(A_1) + \beta_2.R(A_2) + \beta_3.R(A_3). \end{aligned}$$

Conseqüentemente, $g \circ f_i : \{V_{11}, \dots, V_{kn_k}\} \rightarrow R$, onde $g \circ f$ é a composição das duas funções, e se constitui na resposta final tomando por base os dados amostrados e o espaço das variáveis de interesse gerencial: $R = \beta_1.R(A_1) + \beta_2.R(A_2) + \beta_3.R(A_3)$. (3.1)

Existe a possibilidade de desenvolver modelos particulares que representem as combinações das variáveis dos subgrupos, $\{f_1, \dots, f_i\}$, e a superfície de resposta global. No presente trabalho admite-se que as superfícies de respostas são aditivas. LOUVIERE (1994) demonstra outras especificações possíveis que podem ser definidas. Nesta aplicação, são construídos delineamentos experimentais específicos para cada variável latente (área de interesse) e um delineamento experimental especial para a definição da função matemática da composição global.

b) Delineamento experimental especial : GERAL

Delineamento experimental associado às técnicas de preferência declarada para o Modelo Logit Multinomial (LMN), com o objetivo de identificar os coeficientes de importância atribuídos a cada variável latente: *PREÇO*, *QUALIDADE* e *CONVENIÊNCIA*, e estabelecer uma função matemática geral que relaciona as três área em um modelo único para a composição global. A tabela 6.4 apresenta o quadro dos atributos e níveis.

- Função matemática (Mapa de Síntese Geral):

$$R = \beta_1.R(A_1) + \beta_2.R(A_2) + \beta_3.R(A_3) \quad (3.2)$$

Tabela 6.4 – Quadro dos Atributos e os Níveis associados ao Delineamento Experimental Geral

ATRIBUTO	NÍVEIS	Código	FUNÇÃO
(A) Preço	Nível 1: Sem indicação de preferência	0	$\beta_1.R(A_1)$
	Nível 2: Indicação de preferência	1	
(B) Qualidade	Nível 1: Sem indicação de preferência	0	$\beta_2.R(A_2)$
	Nível 2: Indicação de preferência	1	
(C) Conveniência	Nível 1: Sem indicação de preferência	0	$\beta_3.R(A_3)$
	Nível 2: Indicação de preferência	1	

Tabela 6.5 - Formação das Alternativas

Alternativas	Atributos (Códigos)		
	A	B	C
1	1	0	0
2	0	1	0
3	0	0	1

- Os entrevistados serão inquiridos a respeito da ordem de importância que atribuem às variáveis: PREÇO, QUALIDADE e CONVENIÊNCIA, correspondente às alternativas 1, 2 e 3, respectivamente, da tabela 6.5 (processo de ordenamento).

c) Delineamento experimental: PREÇO

Delineamento experimental associado às técnicas de preferência declarada para o Modelo Logit Hierárquico (LH), com o objetivo de identificar a variável latente identificada pelo palavra: *PREÇO* e responder a seguinte questão: *Onde os preços são percebidos pelos clientes?* A tabela 6.6 apresenta o quadro de atributos e níveis associado à área relacionada com a variável latente *PREÇO* e a formação das alternativas consta das tabelas 6.7, 6.8, 6.9, 6.10. Optou-se por LH, tendo em vista que, em testes simulados, os dados não atendiam aos critérios da propriedade IIA.

- Função matemática referente a variável PREÇO:

$$R(A_i) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \quad (3.2)$$

Tabela 6.6 – Quadro dos Atributos e os Níveis associados a PREÇO

ATRIBUTO	NÍVEIS	Código	FUNÇÃO
(A) Carnes	Nível 1: Desconto de 20 %	0	$\beta_1 X_1$
	Nível 2: Desconto de 10 %	1	
	Nível 3: Valor normal sem desconto.	2	
(B) Frutas e verduras	Nível 1: Desconto de 20 %	0	$\beta_2 X_2$
	Nível 2: Desconto de 10 %	1	
	Nível 3: Valor normal sem desconto.	2	
(C) Bens de Mercearia	Nível 1: Desconto de 20 %	0	$\beta_3 X_3$
	Nível 2: Desconto de 10 %	1	
	Nível 3: Valor normal sem desconto.	2	
(D) Laticínios	Nível 1: Desconto de 20 %	0	$\beta_4 X_4$
	Nível 2: Desconto de 10 %	1	
	Nível 3: Valor normal sem desconto.	2	

A realização do experimento se dá em duas etapas:

Primeira etapa: o entrevistado é questionado sobre a preferência pelo conjunto de produtos em que deveria recair um desconto de 20% no preço e, de posse dessa indicação, passa-se para a segunda etapa.

Segunda etapa: em função da escolha efetuada na primeira etapa, passaria a ordenar um dos conjuntos com 3 alternativas formados através da técnica de blocos de cadeia generalizada (Projeto 4.1 – Anexo 3):

- 1) Caso a escolha na primeira etapa seja: CARNES

Tabela 6.7 - Alternativas e níveis dos atributos

Alternativas (Letras)		Atributos (Níveis)			
		A	B	C	D
01	F	*	2	1	0
02	J	*	0	2	1
03	Y	*	1	0	2

- 2) Caso a escolha na primeira etapa seja: FRUTAS E VERDURAS

Tabela 6.8 - Alternativas e níveis dos atributos

Alternativas (Letras)		Atributos (Níveis)			
		A	B	C	D
04	K	0	*	2	1
05	Z	1	*	0	2
06	M	2	*	1	0

3) Caso a escolha na primeira etapa seja: BENS DE MERCEARIAS

Tabela 6.9 - Alternativas e níveis dos atributos

Alternativas (Letras)		Atributos (Níveis)			
		A	B	C	D
07	S	1	0	*	2
08	X	2	1	*	0
09	E	0	2	*	1

4) Caso a escolha na primeira etapa seja: LATICÍNIOS

Tabela 6.10 - Alternativas e níveis dos atributos

Alternativas (Letras)		Atributos (Níveis)			
		A	B	C	D
10	D	2	1	0	*
11	R	0	2	1	*
12	N	1	0	2	*

(*) Indica que o atributo foi escolhido na primeira etapa e, para fins de estimativas dos parâmetros, será interpretado como valor 2.

d) Delineamento experimental: QUALIDADE

Delineamento experimental associado às técnicas de preferência declarada para o modelo Logit Multinomial (LMN), com o objetivo de identificar a variável latente *QUALIDADE* e responder à seguinte questão: *Como é boa a qualidade e a organização?* A tabela 6.11 apresenta o quadro de atributos e níveis associado à área relacionada com a variável latente *QUALIDADE*. Utilizou-se a Tabela A1 – Anexo 2, com a exclusão das alternativas 1 e 8 por serem dominada e dominante, respectivamente, assim o conjunto de escolha ficou composto de 6 alternativas, conforme a tabela 6.12.

- Síntese referente a variável latente *QUALIDADE*:

$$R(A_2) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Figura 6.11 – Quadro dos Atributos e os Níveis associados à Qualidade.

Variável Latente : QUALIDADE / ORGANIZAÇÃO			
Atributos	NÍVEIS	Cod	F U
(A) Carnes	Nível 1: Cortes limitados de carne de boi e galinha. Poucos pacotes de carnes para escolher. Razoavelmente acondicionados.	0	$\beta_1 X_1$
	Nível 2: Todos os tipos de cortes de carne de boi, porco, galinha e carneiro. Sempre com vários pacotes para escolher. Bem acondicionados.	1	
(B) Bens de Mercearia	Nível 1: Produtos regionais (só produtos básicos). Estantes não bem providas e os produtos razoavelmente organizados.	0	$\beta_3 X_3$
	Nível 2: Toda linha nacional e regional. Estantes bem organizadas e bem providas. Estantes de primeira linha.	1	
(C) Frutas e verduras	Nível 1: Itens limitados de frutas e legumes básicos. Sem artigos de especialidades. Sem prateleiras especializadas.	0	$\beta_2 X_2$
	Nível 2: Uma variedade de artigos de especialidade. Prateleiras refrigeradas e verduras sempre frescas.	1	
(D) Laticínios	Nível 1: Alguns produtos básicos. Produtos e escolhas limitadas. Qualidade sem definição. Balcões simples.	0	$\beta_4 X_4$
	Nível 2: Grandes Quantidades de produtos básicos e especiais. Produtos frescos e de primeira qualidade. Balcões com refrigeração especial.	1	

Tabela 6.12 - Códigos Binários das Alternativas

Alternativas (Letras)		Atributos (Códigos)			
		A	B	C	D
01	S	1	1	0	0
02	K	0	0	1	1
03	E	1	0	1	0
04	M	0	1	0	1
05	A	1	0	0	1
06	P	0	1	1	0

e) *Delineamento experimental: CONVENIÊNCIA*

Delineamento experimental associado às técnicas de preferência declarada para o modelo Logit Multinomial com Probabilidade Condicional (LMPC), com o objetivo de identificar a área *CONVENIÊNCIA* e responder à seguinte questão: *Como é conveniente realizar este supermercado?* As tabelas 6.13 e 6.14 apresentam, respectivamente, o quadro de atributos e níveis associado à área relacionada com a variável latente *CONVENIÊNCIA* e seus códigos binários. Adotou-se a tabela A1 do anexo 2.

Tabela 6.13 – Quadro dos Atributos e os Níveis associados à CONVENIÊNCIA

ATRIBUTO	NÍVEIS	Cod	Função
(A) Estacionamento	Nível 1: Limitado e sem segurança	0	$\beta_1 X_1$
	Nível 2: Amplo e com segurança	1	
(B) Lojas de conveniência	Nível 1: Não	0	$\beta_2 X_2$
	Nível 2: Sim (Correios, Farmácia, Revistas, etc.)	1	
(C) Organização Interna	Nível 1: Corredores estreitos e prateleiras razoavelmente organizadas, sem indicações dos produtos.	0	$\beta_3 X_3$
	Nível 2: Corredores largos e produtos bem organizados nas prateleiras e visual alegre e prático para as compras.	1	
(D) Sistema de crédito	Nível 1: Não – venda sempre a vista.	0	$\beta_4 X_4$
	Nível 2: Sim - vários sistemas de créditos	1	
(E) Níveis de Serviços ao Cliente	Nível 1: Nível de serviços simplificado para baratear os produtos. Poucos caixas e funcionários.	0	$\beta_5 X_5$
	Nível 2: Bom nível de serviço. Vários funcionários e muitos caixas disponíveis. Loja sempre bem limpa e higienizada.	1	
(F) Diversificação	Nível 1: Somente os produtos básicos estão disponíveis.	0	$\beta_6 X_6$
	Nível 2: Existe uma diversificação muito grande de produtos disponíveis. (Vende de tudo)	1	
(*) Desconto*	Nível 1: Valor normal dos produtos.	0	$\beta_7 X_7$
	Nível 2: Valor 5% abaixo do valor normal.	1	

(*) Foi acrescentado o atributo Desconto*, para estabelecer um equilíbrio entre as alternativas e que indica o custo médio dos produtos adquiridos naquela loja.

- Síntese referente à variável latente CONVENIÊNCIA:

$$R(A_3) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 (3.4)$$

Tabela 6.14 - Códigos Binários das Alternativas

Alternativas		Atributos (Códigos binários)						
		A	B	C	D	E	F	Desconto*
01	K	0	0	1	1	0	1	1
02	T	0	1	0	1	1	0	1
03	S	0	1	1	0	1	1	0
04	M	1	0	0	1	1	1	0
05	B	1	0	1	0	1	0	1
06	A	1	1	0	0	0	1	1
07	E	1	1	1	1	0	0	0
08	(*)	0	0	0	0	0	0	0

(*) Alternativa com todos os níveis inferiores dos atributos para manter a ortogonalidade do delineamento e evitar combinação linear entre as alternativas ou atributos.

Considerando que a característica das entrevistas seria de abordagem rápida e com pessoas leigas no assunto, houve uma preocupação com o número de alternativas a

serem apresentadas ao entrevistado. Temia-se que um experimento com 7 alternativas e 7 atributos apresentasse grandes erros de percepções e fadiga, então optou-se por dividir o conjunto de 7 alternativas em grupos de 4, utilizando-se da técnica de blocos incompletos balanceados (BIB). A tabela 6.15 apresenta a forma de divisão em 7 blocos ($b = 7$), com 4 alternativas em cada bloco ($k=4$), cada alternativa repete 4 vezes ($r = 4$) e cada alternativa aparece junto com outra alternativa o mesmo número de vezes, $\lambda_1 = 2$.

Tabela 6.15 – Formação do Blocos de Alternativas (Projeto de BIB)

Blocos	Alternativas								Atribuição das alternativas
	01	02	03	04	05	06	07	08 (*)	
01	X	X	X	X				X	1 2 3 4 8
02	X	X			X	X		X	1 2 5 6 8
03	X		X		X		X	X	1 3 5 7 8
04	X			X		X	X	X	1 4 6 7 8
05		X	X			X	X	X	2 3 6 7 8
06		X		X	X		X	X	2 4 5 7 8
07			X	X	X	X		X	3 4 5 6 8

A alternativa número 8, nitidamente inferior a qualquer outra, incluída somente na fase das estimativas dos parâmetros, tem o objetivo de estabelecer a ortogonalidade nos níveis dos atributos. Ela não foi apresentada aos entrevistados e figurou como a última escolha no conjunto de escolha. O experimento apresenta uma eficiência, $E = 0,9143$. Utilizou-se o projeto 4.2 do anexo 4 para a divisão do conjunto de alternativas em conjunto de escolha.

6.2.4. Delineamento experimental com dados de preferência revelada (PR).

Delineamento experimental associado às técnicas de preferência revelada para o Modelo Logit Multinomial (LMN), com o objetivo de estimar os coeficientes da equação de regressão linear múltipla:

$$U^{PR} = \alpha_1 x_{11} + \alpha_2 x_{22} + \alpha_3 x_{33} + \alpha_4 x_{44} + \alpha_5 x_{55} + \varepsilon \quad (3.5)$$

onde:

α é o vetor dos parâmetros e α_i representa a utilidade do Supermercado i ;

X é a matriz de delineamento, $X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ e x_{ij} são os elementos de X .

Cada linha da matriz de delineamento representa um supermercado.

Os parâmetros foram estimados utilizando-se da função de verossimilhança:

$$L^* = \prod_{n=1}^N \prod_{j=1}^2 \frac{e^{\alpha_j x_{ijn}}}{\sum_{i=1}^5 e^{\alpha_i x_{ijn}}}, \text{ onde: } i \text{ indica os atributos (colunas da matriz } X); j$$

representa a indicação de preferência do supermercado e n a entrevista, do Modelo Logit Multinomial e método de maximização de Newton-Raphson, com busca unidirecional por seção áurea.

A tabela 6.16 apresenta o quadro dos atributos e níveis associados ao delineamento experimental da pesquisa de preferência revelada e a tabela 6.17 os atributos relativos às alternativas.

Tabela 6.16 – Atributos e os Níveis associados à Delineamento Experimental

Atributo/Supermercado	NÍVEIS	Código	FUNÇÃO
(A) SM-A1	Nível 1: Sem indicação de preferência	0	$\alpha_1 X_1$
	Nível 2: Indicação de preferência	1	
(B) SM-A2	Nível 1: Sem indicação de preferência	0	$\alpha_2 X_2$
	Nível 2: Indicação de preferência	1	
(C) SM-B1	Nível 1: Sem indicação de preferência	0	$\alpha_3 X_3$
	Nível 2: Indicação de preferência	1	
(D) SM-B2	Nível 1: Sem indicação de preferência	0	$\alpha_4 X_4$
	Nível 2: Indicação de preferência	1	
(E) SM-A3	Nível 1: Sem indicação de preferência	0	$\alpha_5 X_5$
	Nível 2: Indicação de preferência	1	

Tabela 6.17 - Formação das Alternativas

Alternativas	Atributos (Códigos)				
	A	B	C	D	E
01	1	0	0	0	0
02	0	1	0	0	0
03	0	0	1	0	0
04	0	0	0	1	0
05	0	0	0	0	1

Os entrevistados serão estimulados a indicarem dois supermercados onde normalmente realizam suas compras mensais na ordem de preferência.

6.2.5. Integração da preferência declarada com a revelada

A integração entre as utilidades dos supermercados obtida pelos métodos de preferência revelada e declarada será feita segundo o paradigma da figura 6.4.

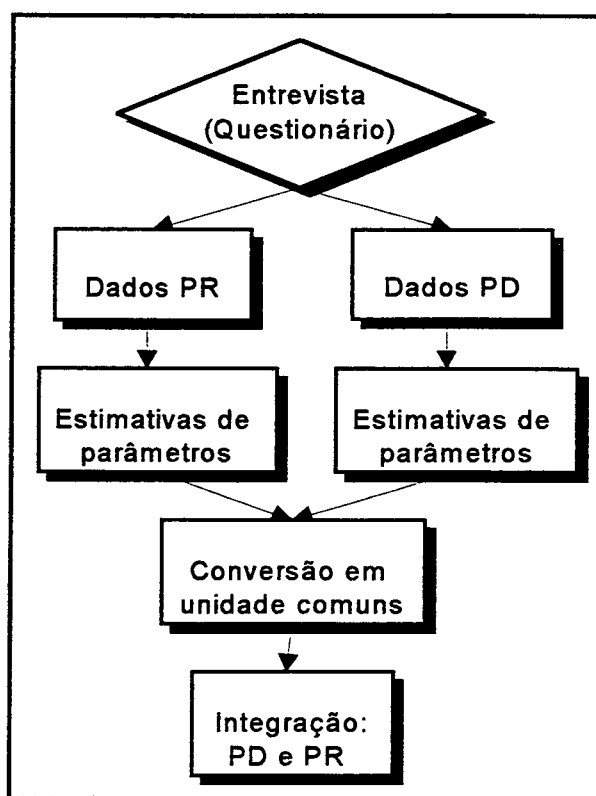


Figura 6.4 – Paradigma de Integração entre PR e PD

Levando-se em consideração que a utilidade real de um dado supermercado possa ser explicada pelo coeficiente de utilidade, U^{PR} , estimado através da equação de

regressão (3.5), então a utilidade, U^{PR} , pode ser escrita como a soma da U^{PD} com um valor, Δ , não explicado pelos atributos da preferência declarada, ou seja: $U^{PR} = U^{PD} + \Delta$. Onde,

$$U^{PR} = \alpha_1 x_{11} + \alpha_2 x_{22} + \alpha_3 x_{33} + \alpha_4 x_{44} + \alpha_5 x_{55} \quad (\text{Preferência Revelada}),$$

$$U^{PD} = \beta_1 R(A_1) + \beta_2 R(A_2) + \beta_3 R(A_3) \quad (\text{Preferência Declarada}).$$

Assim, os valores Δ_i , relativos aos níveis reais dos atributos nos supermercados i ($i = 1, \dots, 5$), pode ser calculado pela diferença vetorial: $\Delta_i = \tilde{U}_i^{PR} - \tilde{U}_i^{PD}$. Os valores das utilidades \tilde{U}_i^{PR} e \tilde{U}_i^{PD} , ($i = 1, \dots, 5$) são as utilidades dos supermercados convertidas em unidades comuns através da proporcionalidade entre os seus valores:

$$\tilde{U}_i^{PR} = \frac{U_i^{PR}}{\sum_{i=1}^5 U_i^{PR}} \quad \text{e} \quad \tilde{U}_i^{PD} = \frac{U_i^{PD}}{\sum_{i=1}^5 U_i^{PD}}.$$

O Coeficiente de Atratividade dos Supermercados (CAS) pode ser obtido pela razão exponencial:

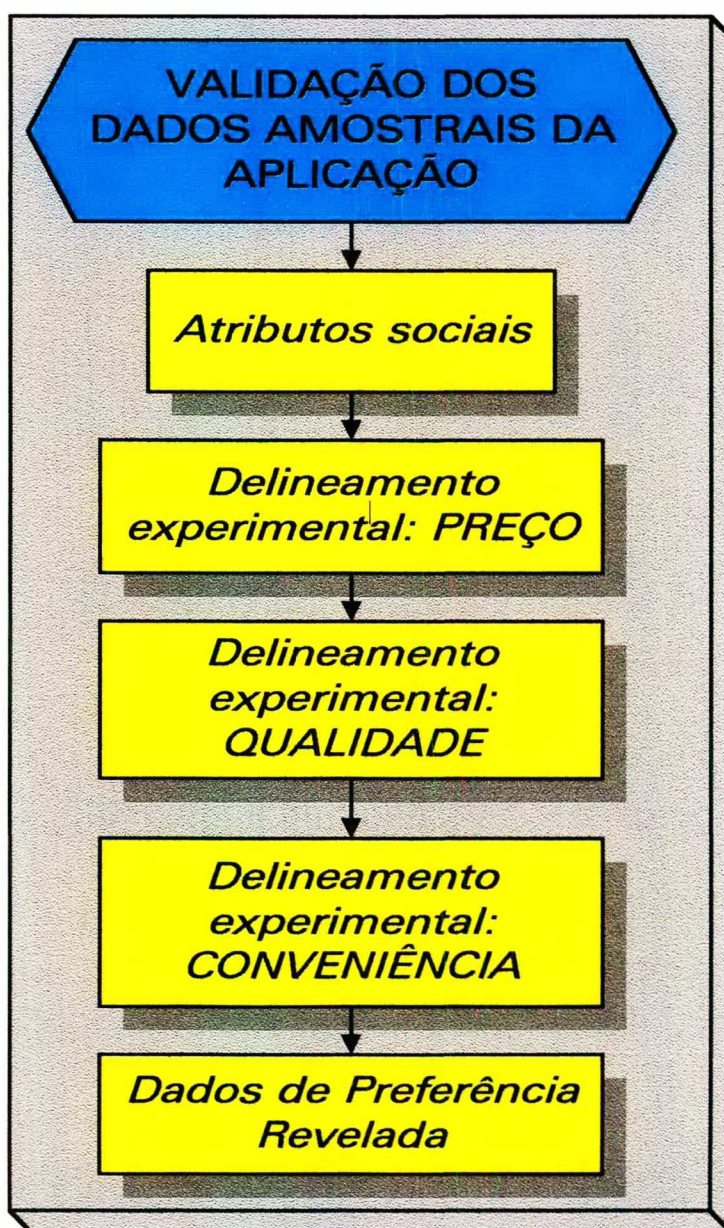
$$CAS_i = \frac{e^{\Delta_i}}{\sum_{i=1}^5 e^{\Delta_i}}.$$

A atratividade dos clientes, por um supermercado, expressa por fatores não definidos nos atributos da pesquisa de preferência declarada.

O CAS representa, matematicamente, o somatório dos fatores não definidos, como atributos, na pesquisa de preferência declarada que geram um diferencial de atratividade para um determinado supermercado.

PARTE II – APLICAÇÃO

Capítulo 7 – VALIDAÇÃO DOS DADOS AMOSTRAIS DA APLICAÇÃO



7. VALIDAÇÃO DOS DADOS AMOSTRAIS DA APLICAÇÃO

7.1. DADOS DESCRITIVOS DOS ATRIBUTOS SOCIAIS

A amostra foi coletada no mês de janeiro de 1999, com cerca de 50% das entrevistas realizadas entre os alunos da UNICENTRO, escolhidos aleatoriamente, e as outras 50% coletadas nos bairros da cidade, também de forma aleatória e sempre reservando uma borda de influência dos supermercados locais, de tal sorte a obter igualdade de representação dos segmentos amostrados. Foram realizadas 175 entrevistas, cujos percentuais referentes à segmentação da amostra (critérios sociais) constam das tabela 7.1 e figura 7.1.

Tabela 7.1 – Segmentação da amostra por atributos sociais

Atributos Sociais	Níveis	Amostra
Sexo	Masculino	80 (46%)
	Feminino	95 (54%)
Escolaridade	1º Grau (completo ou incompletos)	27 (15%)
	2º Grau (completo ou incompletos)	58 (33%)
	3º Grau (completo ou incompletos)	90 (52%)
Idade	Até 20 anos	65 (37%)
	De 21 a 35 a.	97 (56%)
	De 26 a 50 a.	11 (6%)
	Acima de 50 a.	2 (1%)
Renda Familiar	Classe D - Até R\$ 600	39 (22%)
	Classe C - De R\$ 601 a 1200	56 (32%)
	Classe B - De R\$ 1201 a 2400	50 (29%)
	Classe A - Acima de R\$ 2400	30 (17%)

Todas as entrevistas foram executadas pelo autor da tese. O entrevistado, ao ser abordado, era informado a respeito dos objetivos da pesquisa e da metodologia a ser adotada. Seguiu-se o formulário (anexo 7) em todas as entrevistas e o processo de escolha se deu segundo os cenários descritos pelos conjuntos de cartões confeccionados de acordo com os delineamentos experimentais (anexo 8). Todos os cenários retratados nos cartões foram ilustrados com fotografias e desenho característicos. Cerca de 80% dos entrevistados preencheram de próprio punho os formulários, nos outros casos, o

pesquisador foi quem os preencheu, com base nas informações recebidas. Gastou-se em média cerca de 10 minutos por entrevista, em um espaço de 3 dias.

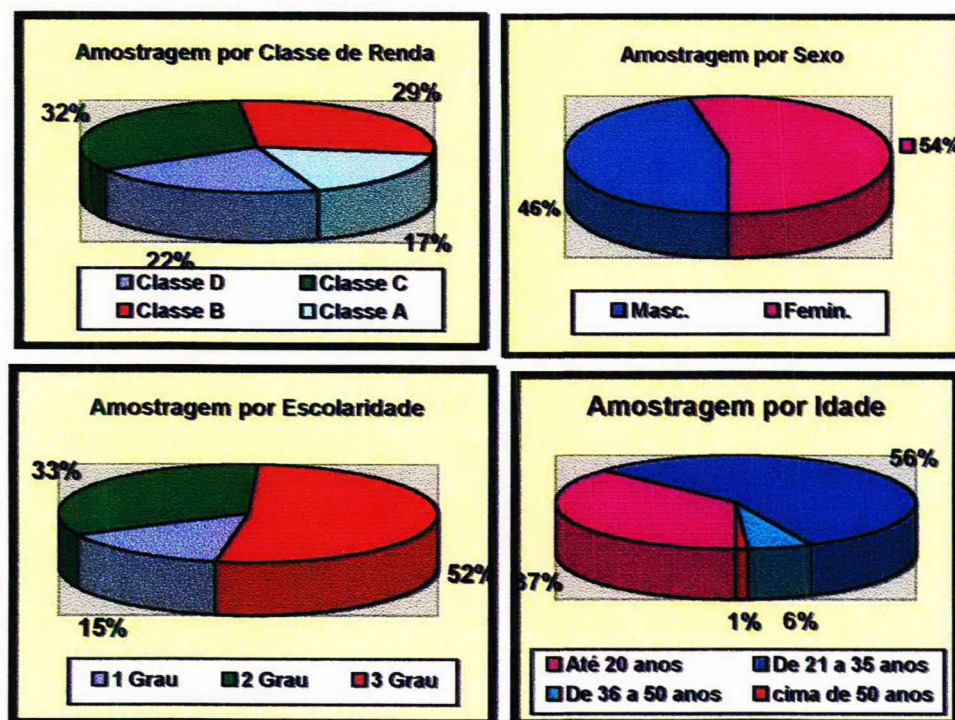


Figura 7.1 – Dados amostrais por atributos sociais

7.2. DADOS DE PREFERÊNCIA DECLARADA: CONDIÇÕES DA AMOSTRA NO DELINEAMENTO EXPERIMENTAL PREÇO

Análise das condições amostrais referentes a possíveis dados discrepantes, presença de heterocedasticidade regular intra-alternativa e dimensionamento da amostra.

7.2.1. Dados discrepantes

O delineamento experimental PREÇO se constitui em um experimento hierárquico, em que a decisão pela preferência de uma alternativa é feita em duas etapas, assim a identificação e a exclusão das alternativas discrepantes alterariam a estrutura do experimento causando viés na relação entre as escolhas, razão pela qual se optou por não adotar tal procedimento.

7.2.2. Transformação dos dados : análise intra-alternativa

Teste de Hartley: $H_{\text{calc}} = 482253189,68$; $H_{\text{tab}} (5\%; 12; 28) = 3,39$;

Resultado: O valor de H_{calc} é superior ao valor H_{tab} , nível 5%, indicando a existência de heterocedasticidade nas variâncias intra-alternativas.

Teste da transformação: $\lambda = -0,69$ $IC(\lambda) = [-350,13 ; 348,74]$; $\rho^2 = 0,1270$;

Resultado: O valor absoluto de λ próximo do valor *um* e o coeficiente de correlação inferior a 0,90 indicam que a correlação entre a média e o desvio padrão nas informações intra-alternativas é indefinida, assim a heterocedasticidade constatada no teste de Hartley é do tipo irregular e provocada pelo tipo de experimento utilizado, Logit Hierárquico, contra-indicando a utilização da transformação dos dados amostrais.

7.2.3. Dimensionamento da amostra

O figura 7.2, apresenta a relação entre o erro (diferença entre os valores dos parâmetros estimados com n amostras e o valores estimados com $n-10$ amostras) e o número de amostras. Para uma estimativa de Erro = 0,05, a função exponencial ajustada indica um dimensionamento de 79 amostras, com o intervalo de confiança [68; 89] e coeficiente de explicação de 98%.

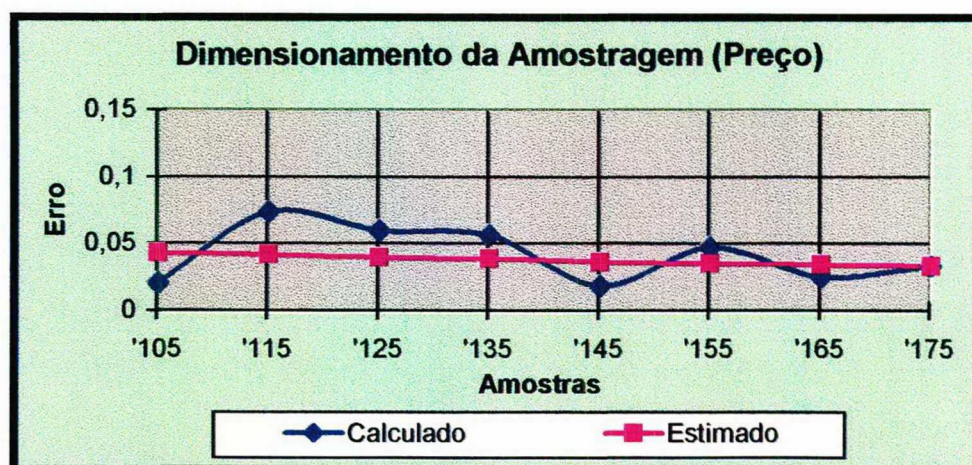


Figura 7.2 – Dimensionamento da amostragem : Preço

Descrição dos resultados: O valor do erro adquire estabilidade para valores abaixo de 0,05 a partir de 137 amostras, indicando que a quantidade de 175 amostras estima assintoticamente valores confiáveis para os parâmetros.

7.3. DADOS DE PREFERÊNCIA DECLARADA: CONDIÇÕES DA AMOSTRA NO DELINEAMENTO EXPERIMENTAL QUALIDADE

Análise das condições amostrais referentes a possíveis dados discrepantes, presença de heterocedasticidade regular intra-alternativa e dimensionamento da amostra:

7.3.1. Dados discrepantes

A tabela 7.2 apresenta os valores utilidades das entrevistas (VUE), agrupadas em intervalos de classe, que apresentou o Coeficiente de Assimetria = -0,807 (Assimetria negativa moderada) que, de acordo com o Teste de Dados Discrepantes (item 4.1.1) indica o percentil discrepante 10 com 21 entrevistas discrepantes, cujos valores utilidade oscilam entre 0,000017 e 0,000542, e 154 entrevistas não discrepantes com valores entre 0,000542 e 0,0072.

Tabela 7.2 – Distribuição das frequências dos Valores Utilidades das Entrevistas

Frequências	Valor Utilidade das entrevistas (Intervalos de classe)						
	0,00001 0,0010	0,0010 0,0021	0,0021 0,0031	0,0031 0,0041	0,0041 0,0051	0,0051 0,0061	0,0061 0,0072
Simples	34 (21*)	10	20	4	21	12	74
Acumulada	34 (21*)	44	64	68	89	101	175

Entrevistas consideradas discrepantes (21*): 7; 8; 11; 27; 30; 34; 37; 50; 87; 89; 93; 99; 100; 112; 113; 116; 132; 135; 139; 142; 155 (que serão excluídas) e a análise será efetuada em 154 entrevistas. As estimativas, pós exclusão dos dados discrepantes, indicaram uma melhoria de performance de 11%.

7.3.2. Transformação dos dados : análise intra-alternativas

A tabela 7.3 apresenta os resultados dos testes de Hartley e da transformação para análise e identificação da necessidade ou não da utilização da transformação logarítmica, pós a exclusão dos dados discrepantes.

Tabela 7.3 – Quadro comparativo entre os Testes de Hartley e transformação

Testes	Dados originais	Dados transformados
Hartley (H_{calc})	87,54 ($H_{\text{tab}} (5\%, 6; 144) = 1,00$)	1,61 ($H_{\text{tab}} (5\%, 6; 144) = 1,00$)
Transformação (λ)	-0,05 [-0,055; -0,045]	0,89 [0,752 ; 1,035]
Correlação (ρ^2)	0,9984	0,3705
Resultado	Heterocedasticidade regular (*)	Homocedasticidade (**)

(*) O valor H_{calc} maior que $H_{\text{tab}} (5\%, 6; 144)$ indica que existe heterocedasticidade intra-alternativas; o valor de λ próximo a *zero* e o coeficiente de correlação maior do que 0,90 indicam que a heterocedasticidade apontada pelo teste de Hartley é do tipo *regular* e pode ser corrigida pela utilização da transformação logarítmica.

(**) Os resultados dos testes aplicados aos dados transformados (Logaritmo), indicam que, praticamente, eliminou-se a heterocedasticidade regular intra-alternativas; testes de Hartley bem próximos ao valor tabelado e o teste da transformação superior a 0,5 com correlação inferior a 0,90. Resultado: rejeita-se a hipótese da existência heterocedasticidade regular das variâncias intra-alternativas, isto significa que a transformação surtiu efeito, conforme pode ser observado nos figura 7.3 e 7.4.

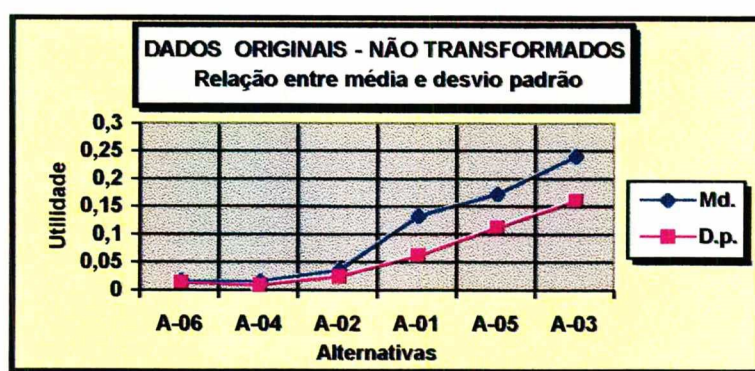


Figura 7.3 – Relação entre a média e o desvio padrão – dados não transformados

Descrição dos resultados: Percebe-se que existe relação entre a média e o desvio padrão das probabilidades de utilidades marginais das alternativas, $P_n (A)_k$, item 4.3, o que caracteriza a heterocedasticidade regular.

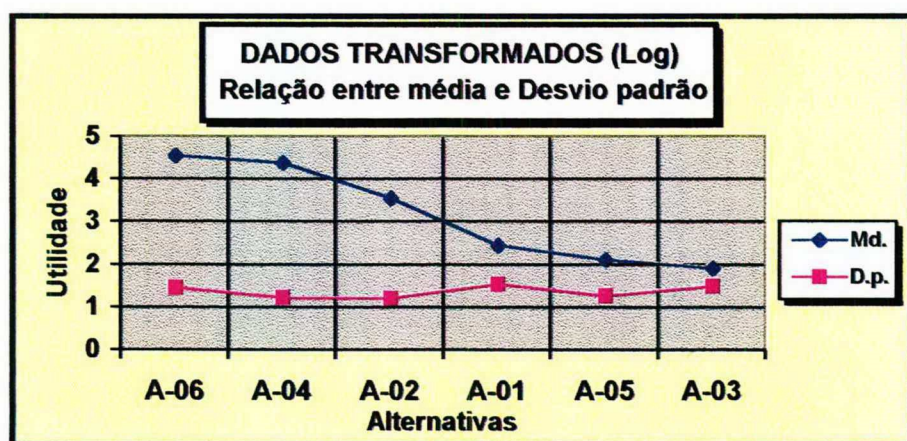


Figura 7.4 – Relação entre a média e o desvio padrão - transformados logarítmica

Descrição dos resultados: Percebe-se que não existe relação entre a média e o desvio padrão das probabilidades de utilidades marginais das alternativas, $P_n(A)_k$, item 4.3, o que caracteriza a homocedasticidade.

7.3.3. Dimensionamento da amostra

O figura 7.5 apresenta a relação entre o Erro (diferença das estimativas dos parâmetros com n amostras e as estimativas com $n-10$ amostras) e o número de amostras. Para uma estimativa de erro = 0,05 a função exponencial ajustada indica um dimensionamento de 118 amostras, com o intervalo de confiança (109; 126) e coeficiente de explicação de 98%.

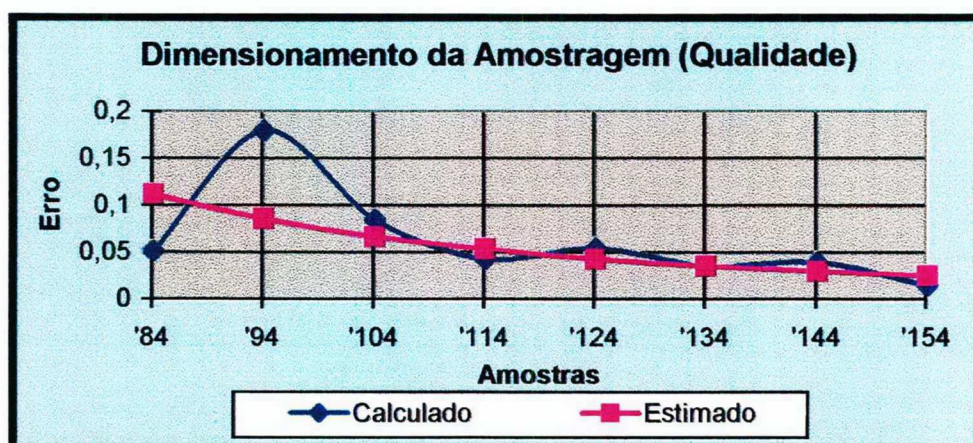


Figura 7.5 – Dimensionamento da Amostragem : Qualidade
(Amostragem sem os dados discrepantes)

Descrição dos resultados: O valor do erro adquire estabilidade para valores abaixo de 0,05 a partir de 130 amostras, indicando que a quantidade de 154 amostras estima assintoticamente valores confiáveis para os parâmetros.

7.4. DADOS DE PREFERÊNCIA DECLARADA: CONDIÇÕES DA AMOSTRA NO DELINEAMENTO EXPERIMENTAL CONVENIÊNCIA

Análise das condições amostrais referentes a possíveis dados discrepantes, presença de heterocedasticidade regular intra-alternativa e dimensionamento da amostra:

7.4.1. Dados discrepantes

O delineamento experimental CONVENIÊNCIA se constitui em um experimento em que o conjunto de escolha foi dividido em subconjuntos menores, utilizando-se das técnicas de BIB. Esse tipo de delineamento experimental apresenta a característica da escolha condicional, sendo que a primeira escolha, ou seja, a escolha do bloco, não é realizada pelo entrevistado; assim, impõe-se a condição da igualdade de representação dos blocos como um pressuposto fundamental; caso contrário, estabelecer-se-ia um modelo hierárquico em que a hierarquia seria definida pelas diferenças de representatividade entre os blocos. Dessa forma é conveniente adotar um dos procedimentos: substituir as entrevistas discrepantes por outras do mesmo conjunto de escolha ou optar pela não exclusão das alternativas discrepantes. Neste experimento, o conjunto dos VUE tende para a distribuição normal com coeficiente de assimetria = -0,075 (Tipo: considerada simétrica) e de acordo com o teste TDD (item 4.1), não indicam a necessidade de exclusão de alternativas consideradas discrepantes.

7.4.2. Transformação dos dados : análise intra-alternativa

Dados originais

Teste de Hartley: $H_{\text{calc}} = 1,61$; $H_{\text{tab}}(5\%; 7; 74) = 2,17$;

Resultado: O valor de H_{calc} é inferior ao valor H_{tab} , nível 5%, indicando a existência de homocedasticidade nas variâncias intra-alternativas;

Teste da transformação: $\lambda = 1,07$ $IC(\lambda) = [0,96 ; 1,18]$; $\rho^2 = -0,2684$;

Resultado: O valor de λ próximo do valor 1 e o coeficiente de correlação inferior a 0,90 indicam que não existe correlação entre a média e o desvio padrão nas informações intra-alternativas, o que corrobora com o teste de Hartley na rejeição da utilização da transformação logarítmica.

7.4.3. Dimensionamento da amostra

O figura 7.6 apresenta a relação entre o Erro (diferença das estimativas dos parâmetros com n amostras e as estimativas com $n-10$ amostras) e o número de amostras. Para uma estimativa de erro = 0,05 a função exponencial ajustada indica um dimensionamento de 174 amostras, com o intervalo de confiança (167; 181) e coeficiente de explicação de 98%.

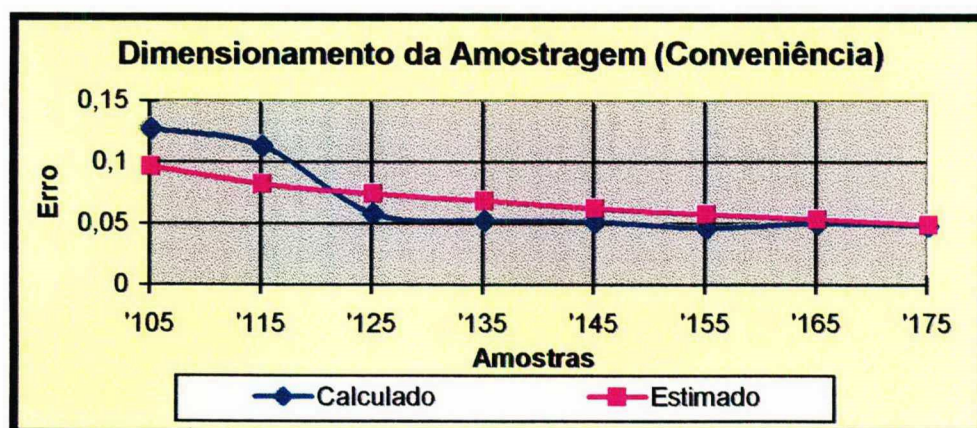


Figura 7.6 – Dimensionamento da Amostragem : Conveniência

Descrição dos resultados: O valor do erro adquire estabilidade, para valores próximos de 0,05, a partir de 135 amostras, indicando que a quantidade de 175 amostras estima assintoticamente valores confiáveis para os parâmetros.

7.4.4. Verificação do ponto de máximo da função

A tabela 7.4 apresenta os resultados obtidos na verificação da convergência do ponto de máximo da função, quando os valores iniciais dos parâmetros são diferentes de zero. Conforme o critério da região de confiança (CRC), item 4.2.1., as estimativas obtidas com valores diferentes de zero não diferem, estatisticamente, dos valores estimados partindo de zero.

Tabela 7.4 – Verificação do ponto de máximo da função por aproximação

Valores iniciais	Parâmetros estimados para o ponto de máximo da FU do delineamento : CONVENIÊNCIAS								Teste da Região de Confiança (*)
0	0,981	0,677	0,971	1,109	0,970	0,800	0,665		-x-
0,3	0,941	0,638	0,932	1,070	0,931	0,761	0,626		0,1227
0,5	1,004	0,689	0,974	1,123	0,978	0,819	0,674		0,0393
0,7	1,078	0,759	1,028	1,185	1,039	0,891	0,738		0,5856
-0,5	1,033	0,729	1,018	1,156	1,015	0,853	0,719		0,2023

Valores tabelados: Limite de aceitação (95%) : 14,1 e Limite Excelente (5%) : 2,17.

(*) RC : De acordo com o teste de Qui-quadrado, com gl. = 7, *ACEITA-SE* a hipótese de que os parâmetros estimados com valores diferentes de 0 sejam incluídos na região de confiança dos parâmetros estimados a partir de 0, *SEM* restrições.

7.5. DADOS DE PREFERÊNCIA REVELADA: CONDIÇÕES DA AMOSTRA.

A tabela 7.5 apresenta a tabulação dos dados da indicação de dois supermercados, na ordem de preferência, onde os entrevistados realizam suas compras mensais. A composição da lista utilizada para estímulo da indicação de preferências obedeceu aos seguintes critérios: (a) composição estritamente didática; (b) os maiores estabelecimentos comerciais do ramo; (c) apresentam características marcantes quanto aos tipos classificados na tabela 6.1. Tendo em vista os objetivos propostos, a tabela 7.6 apresenta os dados segmentados por renda, e a figura 7.7 por tipo de mercado.

Tabela 7.5 - Indicação de Preferências por Compras Mensais

Renda Familiar	SM-A1		SM-A2		SM-A3		SM-B1		SM-B2		Total
	1 ^o	2 ^o	1 ^o	2 ^o	1 ^o	2 ^o	1 ^o	2 ^o	1 ^o	2 ^o	
Classe D	3	4	0	6	1	6	33	2	2	21	78
Classe C	2	13	8	22	9	2	29	18	8	1	112
Classe B	16	8	2	8	9	27	19	7	4	0	100
Classe A	10	2	12	11	0	8	4	9	4	0	60
Subtotal	31	27	22	47	19	43	85	36	18	22	350
Total	58		69		61		121		40		350

Tabela 7.6 - Indicação de Preferências por Tipo de Supermercado

Renda Familiar	Tipo A		Tipo B		Total
	1 ^o	2 ^o	1 ^o	2 ^o	
Classe D	4	16	35	23	78
Classe C	19	37	37	19	112
Classe B	27	43	23	7	100
Classe A	22	21	8	9	60
Subtotal	72	117	103	58	350
Total	189		161		350

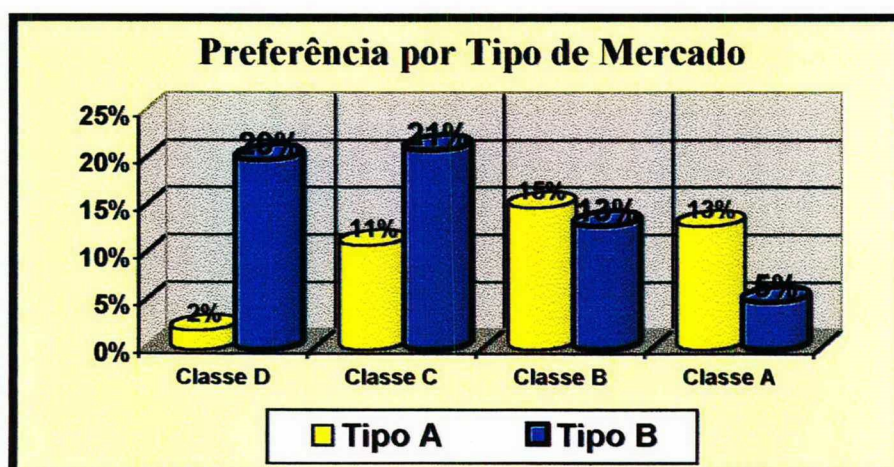
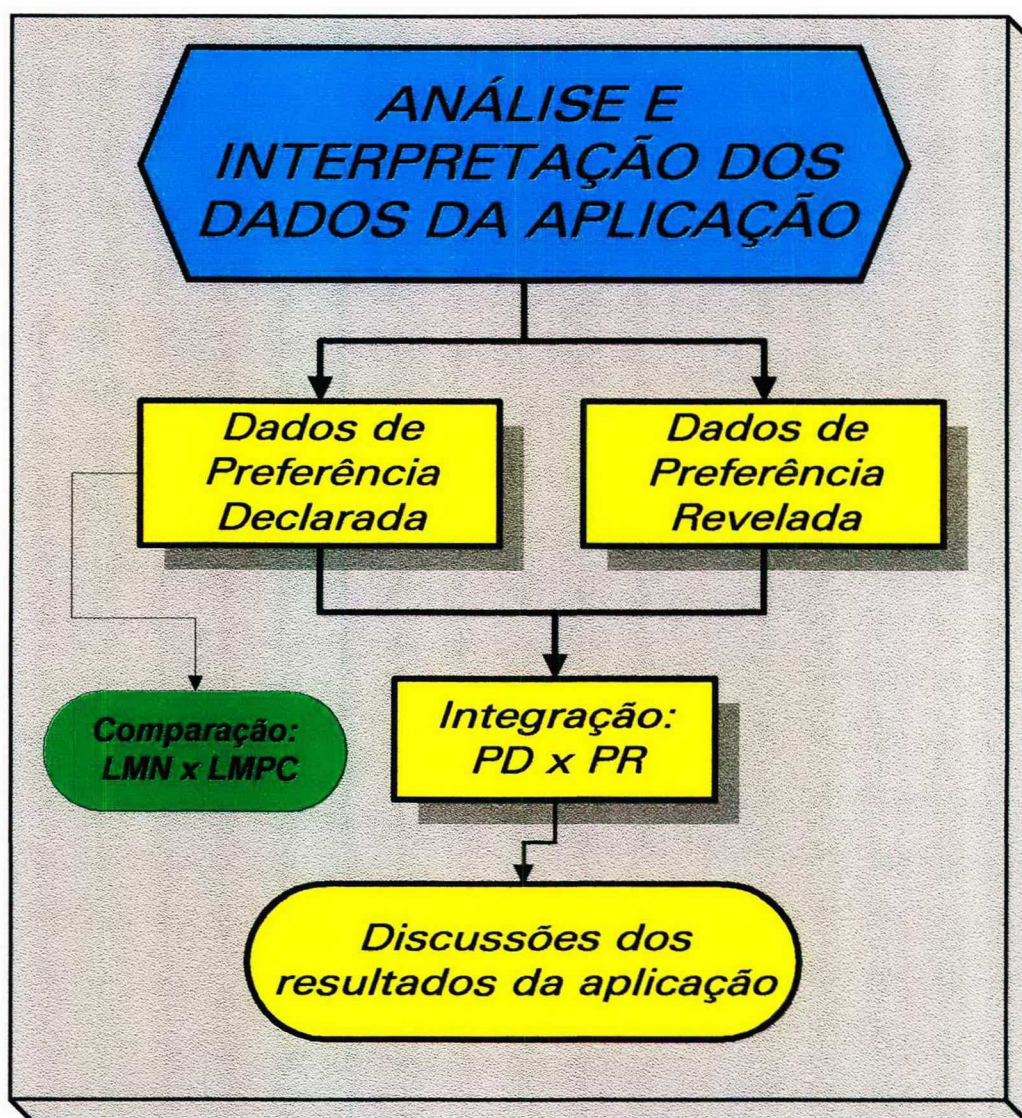


Figura 7.7 – Indicação de Preferências : por Renda

Descrição dos resultados: Percebe-se, na exposição gráfica, que existe um diferenciação nítida no comportamento dos clientes, quando classificados por classes de renda e tipo de mercado. A Classe D (renda até R\$ 600) indica preferência nítida pelo mercado Tipo B (com níveis de serviços simplificados para baratear o custo dos produtos) e, à medida que a renda aumenta, as posições vão sendo modificadas até atingir uma situação oposta.

PARTE II – APLICAÇÃO

Capítulo 8 – ANÁLISE E INTERPRETAÇÃO DOS DADOS DA APLICAÇÃO



8. ANÁLISE E INTERPRETAÇÃO DA APLICAÇÃO

8.1. DADOS DE PREFERÊNCIA DECLARADA

8.1.1. Validação interna: testes de hipóteses

a) *Delineamento experimental: GERAL*

Para a estimativa dos parâmetros, utilizou-se o método de Newton-Raphson, com busca unidirecional por seção áurea. Adotou-se o Modelo Logit Multinomial, tendo em vista que as alternativas formavam um único conjunto de escolha e o delineamento atende aos critérios das propriedades IID (Identicamente e independentemente Gumbel distribuído) e IIA (Independência das alternativas irrelevantes) - no experimento não há alternativa irrelevante. A tabela 8.1 apresenta as estimativas dos coeficientes obtidas com 4 iterações, a partir dos dados coletados (Anexo 9).

Tabela 8.1 – Estimativas dos Coeficientes (GERAL)

Atributos	Coeficiente	Erro	Teste t	IC.(t=2,5%)
Conveniência	1,8807	0,3412	5,5119	[1,21 ; 2,55]
Preço	2,4405	0,3317	7,3575	[1,79 ; 3,09]
Qualidade	4,2124	0,3528	11,9401	[3,52 ; 4,90]

Transformação: DADOS NÃO TRANSFORMADOS

Número de Entrevistas = 175

Número de Casos = 350

$\mathcal{L}(0) = -434,8587$

$\mathcal{L}(\beta) = -225,9217$

$LR (-2[\mathcal{L}(0) - \mathcal{L}(\beta)]) = 417,8739$

$\rho^2 = 0,4805$

$\rho^2_{A_j} = 0,4736$

Descrição dos resultados: Conforme os testes t, LR e ρ^2 , os valores obtidos pelos coeficientes (β) são significativos e assintoticamente convergentes para o máximo da função. Os valores positivos dos coeficientes indicam um aumento da utilidade à medida que um atributo passa do nível 0 para o nível 1.

Os coeficientes estimados serão utilizados como fatores de proporcionalidade de importância na construção da Função Utilidade Geral, ou seja: $R = \beta_1.R(A_1) + \beta_2.R(A_2) + \beta_3.R(A_3)$. Substituindo os valores, a FU geral fica:

$$R = 2,4405.R(A_1) + 4,2124.R(A_2) + 1,8807.R(A_3). \quad (8.1)$$

b) Delineamento experimental: PREÇO

Para a estimativa dos parâmetros, utilizou-se o método de Newton-Raphson, com busca unidirecional por seção áurea. Adotou-se o Modelo Logit Hierárquico com o processo de escolha em duas etapas: primeiro o entrevistado indica um atributo e em seguida passa a ordenar as alternativas de acordo com o atributo escolhido. A razão pela qual se adotou o modelo Logit Hierárquico foi que os testes da propriedade IIA, no modelo Logit Multinomial, deram significativos, em exercícios simulados, indicando a necessidade do modelo LH. A propriedade IID ficou preservada no delineamento adotado. Uma das imposições do modelo LH, é que o cálculo das estimativas dos coeficientes deve levar em conta a probabilidade condicional. Para tanto, adotou-se o modelo denominado LMPC, desenvolvido nesta tese, que leva em conta a probabilidade condicional. A tabela 8.2 apresenta as estimativas dos coeficientes obtidas com 5 iterações.

Tabela 8.2 – Estimativas dos Coeficientes (PREÇO)

Atributos	Coeficiente	Erro	Teste t	IC.(t=2,5%)
Carnes	1,0358	0,1214	8,5300	[0,793 ; 1,279]
Frutas e Verduras	0,3374	0,1213	2,7828	[0,095 ; 0,580]
Bens de Mercearia	0,6067	0,1192	5,0889	[0,368 ; 0,845]
Frios e Laticínios	0,8487	0,1195	7,1052	[0,610 ; 1,088]

Transformação: DADOS NÃO TRANSFORMADOS

Número de Entrevistas = 175

Número de Casos = 350

$\mathcal{L}(0) = -434,8587$

$\mathcal{L}(\beta) = -309,0548$

$LR (-2[\mathcal{L}(0) - \mathcal{L}(\beta)]) = 251,6078$

$\rho^2 = 0,2893$

$\rho^2_{Aj} = 0,2801$

Descrição dos resultados: Conforme os testes t, LR e ρ^2 , os valores obtidos pelos coeficientes (β) são significativos e assintoticamente convergentes para o máximo da função, corroborado com os resultados obtidos no dimensionamento da amostra. Os valores positivos dos coeficientes indicam um aumento da utilidade à medida que um atributo passa do nível inferior para um nível superior.

Função Utilidade geral definida no delineamento PREÇO:

$$R(A_1) = \beta_1.X_{11} + \beta_2.X_{12} + \beta_3.X_{13} + \beta_4.X_{14}$$

$$R(A_1) = 1,0358.X_{11} + 0,3374.X_{12} + 0,6067.X_{13} + 0,8487.X_{14} \quad (8.2)$$

c) *Delineamento experimental: QUALIDADE*

Para a estimativa dos parâmetros, utilizou-se o método de Newton-Raphson, com busca unidirecional por seção áurea. Adotou-se o Modelo Logit Multinomial (LMN) com o processo de escolha explodida (ordenação). As propriedades IID e IIA foram preservadas no delineamento adotado. A tabela 8.3 apresenta as estimativas dos coeficientes obtidas com 4 iterações. Utilizou-se a transformação logarítmica de acordo com a indicação descrita no item 7.3 e foram excluídas 21 alternativas consideradas discrepantes.

Tabela 8.3 – Estimativas dos Coeficientes (QUALIDADE)

Atributos	Coeficiente	Erro	Teste t	IC.(t=2,5%)
Carnes	4,4651	0,5043	8,8536	[3,48 ; 5,45]
Mercearia	1,7489	0,3863	4,5275	[0,99 ; 2,51]
Frutas e Verduras	0,7539	0,3859	1,9535	[0,00 ; 1,51]
Laticínios	1,5360	0,3820	4,0205	[0,79 ; 2,28]

Transformação utilizada : LOGARITMO NEPERIANO

Número de Entrevistas = 154

Número de Casos = 770

$\mathcal{L}(0) = -316,9658$

$\mathcal{L}(\beta) = -227,7783$

$LR (-2[\mathcal{L}(0) - \mathcal{L}(\beta)]) = 178,3752$

$\rho^2 = 0,2814$

$\rho^2_{Aj} = 0,2688$

Descrição dos resultados: Conforme os testes t, LR e ρ^2 , os valores obtidos pelos coeficientes (β) são significativos e assintoticamente convergentes para o máximo da função, corroborado com os resultados obtidos no dimensionamento da amostra. Os valores positivos dos coeficientes indicam um aumento da utilidade à medida que um atributo passa do nível 0 para o nível 1.

Função Utilidade geral definida no delineamento PREÇO:

$$R(A_2) = \beta_1.X_{21} + \beta_2.X_{22} + \beta_3.X_{23} + \beta_4.X_{24}$$

$$R(A_2) = 4,4651.X_{21} + 1,7489.X_{22} + 0,7539.X_{23} + 1,5360.X_{24} \quad (8.3)$$

d) *Delineamento experimental: CONVENIÊNCIA*

Para a estimativa dos parâmetros, utilizou-se o método de Newton-Raphson, com busca unidirecional por seção áurea. Adotou-se o Modelo Logit Multinomial com Probabilidade Condicional, desenvolvido nesta tese, e o processo de escolha explodida nos blocos (ordenação). A razão pela qual se adotou o modelo LMPC foi que o conjunto

de escolha continha 8 alternativa, quantidade considerada muito grande para submeter ao entrevistado; assim, o conjunto de escolha foi dividido em 7 blocos com 4 alternativas cada, impondo a necessidade de se levar em conta a probabilidade condicional e o peso dos blocos nas estimativas dos coeficientes e garantir que as alternativas sejam IID. A propriedade IIA ficou preservada no delineamento adotado. A tabela 8.4 apresenta as estimativas dos coeficientes obtidas com 5 iterações.

Tabela 8.4 – Estimativas dos Coeficientes (CONVENIÊNCIA)

Atributos	Coeficiente	Erro	Teste t	IC.(t=2,5%)
Estacionamento	0,9815	0,1445	6,7938	[0,693 ; 1,270]
Lojas de Conveniência	0,6775	0,1433	4,7269	[0,391 ; 0,964]
Organização	0,9713	0,1439	6,7490	[0,683 ; 1,259]
Sistema crédito	1,1096	0,1425	7,7853	[0,825 ; 1,395]
Níveis serviços	0,9700	0,1448	6,6988	[0,680 ; 1,260]
Diversificação	0,8009	0,1441	5,5566	[0,513 ; 1,089]
Custo 5% abaixo da média	0,6659	0,1449	4,5962	[0,376 ; 0,956]

Eficiência = 0,9143 (A eficiência já está incluída na Variância.)

Transformação: DADOS NÃO TRANSFORMADOS

Número de Entrevistas = 175

Número de Casos = 525

$\mathcal{L}(0) = -716,5103$

$\mathcal{L}(\beta) = -550,9847$

$LR (-2[\mathcal{L}(0) - \mathcal{L}(\beta)]) = 331,0511$

$\rho^2 = 0,2310$

$\rho^2_{Aj} = 0,2212$

Descrição dos resultados: Conforme os testes t, LR e ρ^2 , os valores obtidos pelos coeficientes (β) são significativos e assintoticamente convergentes para o máximo da função, corroborado com os resultados obtidos no dimensionamento da amostra. Os valores positivos dos coeficientes indicam um aumento da utilidade à medida que um atributo passa do nível 0 para o nível 1.

Função Utilidade geral definida no delineamento PREÇO:

$$R(A_3) = \beta_1.X_{31} + \beta_2.X_{32} + \beta_3.X_{33} + \beta_4.X_{34} + \beta_5.X_{35} + \beta_6.X_{36} + \beta_7.X_{37}$$

$$R(A_3) = 0,9815.X_{31} + 0,6775.X_{32} + 0,9713.X_{33} + 1,1096.X_{34} + 0,9700.X_{32} + \\ + 0,8009.X_{33} + 0,6659.X_{34} \quad (8.4)$$

8.1.2. Construção da Função Utilidade

A Função Utilidade expressa a visão da população amostrada com referência aos atributos apresentados e tem por finalidade representar matematicamente a síntese geral

da visão dos clientes. Constitui-se em uma superfície de resposta em que é possível identificar, sendo conhecidos os atributos reais de um dado Supermercado, sua utilidade (atributos fortes e fracos) do ponto de vista da população.

a) Função Utilidade (Geral)

A Função Utilidade Geral é dada pela composição das função $R(A_1)$, $R(A_2)$, $R(A_3)$ em R:

$$R_{(Geral)} = 2,4405.R(A_1) + 4,2124.R(A_2) + 1,8807.R(A_3). \quad (8.5)$$

$$R(A_1) = 1,0358.X_{11} + 0,3374.X_{12} + 0,6067.X_{13} + 0,8487.X_{14}$$

$$R(A_2) = 4,4651.X_{21} + 1,7489.X_{22} + 0,7539.X_{23} + 1,5360.X_{24}$$

$$R(A_3) = 0,9815.X_{31} + 0,6775.X_{32} + 0,9713.X_{33} + 1,1096.X_{34} + 0,9700.X_{32} + \\ + 0,8009.X_{33} + 0,6659.X_{34}$$

b) Função Utilidade (Classe D - Renda Familiar até R\$ 600)

A Função Utilidade segmentada para renda familiar até R\$ 600 é formada pelos coeficientes estimados a partir da segmentação da amostra e dada pela composição das função $R(A_1)$, $R(A_2)$, $R(A_3)$, em R:

$$R_{(Classe A)} = 2,04.R(A_1) + 4,98.R(A_2) + 2,14.R(A_3). \quad (8.6)$$

$$R(A_1) = 0,9377.X_{11} + 0,7674.X_{12} + 0,3384.X_{13} + 0,8672.X_{14}$$

$$R(A_2) = 3,9132.X_{21} + 0,8288.X_{22} + 1,7090.X_{23} + 1,3900.X_{24}$$

$$R(A_3) = 0,6765.X_{31} + 0,5352.X_{32} + 0,8594.X_{33} + 1,4460.X_{34} + 1,3418.X_{32} + \\ + 0,8117.X_{33} + 0,5741.X_{34}$$

c) Função Utilidade (Classe C - Renda Familiar de R\$ 601 a 1200)

A Função Utilidade segmentada para renda familiar de R\$ 601 a 1200 é formada pelos coeficientes estimados a partir da segmentação da amostra e dada pela composição das função $R(A_1)$, $R(A_2)$, $R(A_3)$, em R:

$$R_{(Classe B)} = 2,36.R(A_1) + 3,62.R(A_2) + 2,18.R(A_3). \quad (8.7)$$

$$R(A_1) = 1,1851.X_{11} + 0,3529.X_{12} + 0,3984.X_{13} + 0,9125.X_{14}$$

$$R(A_2) = 3,8421.X_{21} + 1,1750.X_{22} + 1,6561.X_{23} + 1,0109.X_{24}$$

$$R(A_3) = 1,1679.X_{31} + 0,5370.X_{32} + 1,2908.X_{33} + 1,0416.X_{34} + 0,7122.X_{32} + \\ + 1,0203.X_{33} + 0,4628.X_{34}$$

d) Função Utilidade (Classe B - Renda Familiar de R\$ 1201 a 2400)

A Função Utilidade segmentada para renda familiar de R\$ 1201 a 2400 é formada pelos coeficientes estimados a partir da segmentação da amostra e dada pela composição das função $R(A_1)$, $R(A_2)$, $R(A_3)$, em R:

$$R_{(Classe C)} = 2,71.R(A_1) + 4,49.R(A_2) + 1,38.R(A_3). \quad (8.8)$$

$$R(A_1) = 0,8448.X_{11} + 0,2012.X_{12} + 0,9546.X_{13} + 0,7922.X_{14}$$

$$R(A_2) = 5,1931.X_{21} + 0,7140.X_{22} + 1,6179.X_{23} + 1,8819.X_{24}$$

$$R(A_3) = 0,7063.X_{31} + 0,8762.X_{32} + 1,0219.X_{33} + 0,9748.X_{34} + 1,2335.X_{32} + 0,5277.X_{33} + 0,8510.X_{34}$$

e) Função Utilidade (Classe A - Renda Familiar acima de R\$ 2400)

A Função Utilidade segmentada para renda familiar acima de R\$ 2400 é formada pelos coeficientes estimados a partir da segmentação da amostra e dada pela composição das função $R(A_1)$, $R(A_2)$, $R(A_3)$, em R:

$$R_{(Classe D)} = 2,64.R(A_1) + 4,83.R(A_2) + 1,25.R(A_3). \quad (8.9)$$

$$R(A_1) = 1,0473.X_{11} + -0,4940.X_{12} + 0,5460.X_{13} + 0,4723.X_{14}$$

$$R(A_2) = 7,7810.X_{21} + -0,5143.X_{22} + 2,4826.X_{23} + 1,8512.X_{24}$$

$$R(A_3) = 1,4524.X_{31} + 0,7996.X_{32} + 0,5567.X_{33} + 1,1559.X_{34} + 0,6196.X_{32} + 0,8216.X_{33} + 0,7960.X_{34}$$

8.1.3. Análise Comparativa entre LMPC e LMN

a) Delineamento Experimental : PREÇO

Com o objetivo de estabelecer uma comparação entre os coeficientes estimados, levando-se em conta probabilidade condicional da escolha hierárquica (Modelo Logit Multinomial com Probabilidade Condicional – LMPC) e os coeficientes estimados sem levar em conta a probabilidade condicional (Modelo Logit Multinomial – LMN), procedeu-se às estimativas nos dois modelos (tabela 8.5) e utilizaram-se os testes da RC, HM e MTT para diagnosticar a comparação dos dados do delineamento experimental denominado PREÇO (Logit Hierárquico).

Tabela 8.5 – Estimativa dos coeficientes : PREÇO

Atributos	LMPC		LMN	
	Coeficiente	Teste t	Coeficiente	Teste t
Carnes	1,0358	8,5300	0,6812	5,3264
Frutas e Verduras	0,3374	2,7828	0,5054	4,1477
Mercearia	0,6067	5,0889	0,7125	5,8560
Laticínios	0,8487	7,1052	1,0344	8,5703
$\mathcal{L}(0)$ e $\mathcal{L}(\beta)$	-434,8587	-309,0548	-434,8587	-298,7534
LR	251,6078	-	272,2105	-
ρ^2	0,2893	-	0,3130	-

Delineamento Experimental : PREÇO (Matrizes de Covariâncias)

LMPC				LMN			
0,0147	0,0088	0,0091	0,0093	0,0164	0,0077	0,0075	0,0078
0,0088	0,0147	0,0106	0,0107	0,0077	0,0149	0,0104	0,0105
0,0091	0,0106	0,0142	0,0102	0,0075	0,0104	0,0148	0,0103
0,0093	0,0107	0,0102	0,0143	0,0078	0,0105	0,0103	0,0146

Valor do $\chi^2_{(4)}$: Limite Excelente (5%) : 0,711; Limite de aceitação (95%) : 9,49

Teste da Razão de Confiança: RC = 31,59 ;

Resultado: => Os coeficientes estimado por LMN estão fora do IC dos coeficientes estimados por LMPC.

Teste de Comparação: HM = 91,31; MTT = 20,60.

Resultado: => Rejeita-se a hipótese de igualdade dos coeficientes.

Descrição dos resultados: Nesse caso, os coeficientes estimados são estatisticamente diferentes e refletem informações claras de que em experimentos LH não se recomenda a utilização do Modelo LMN (DALY, 1987); para as estimativas há necessidade de se levar em conta a probabilidade condicional da 1ª escolha, caso contrário não estaria atendendo à propriedade IID e os resultados estariam viesados. A variância em LMPC é menor do que em LMN. BEN-AKIVA e BOCCARA (1995) também obteve conclusões semelhantes.

b) Delineamento Experimental : CONVENIÊNCIA

Com o objetivo de estabelecer uma comparação entre os coeficientes estimados, levando-se em conta probabilidade condicional da escolha hierárquica (Modelo Logit Multinomial com Probabilidade Condicional – LMPC) e os coeficientes estimados sem levar em conta a probabilidade condicional (Modelo Logit Multinomial – LMN),

procedeu-se às estimativas nos dois modelos (tabela 8.6) e utilizaram-se os testes da RC, HM e MTT para diagnosticar a comparação dos dados do delineamento experimental denominado CONVENIÊNCIA (Logit com o conjunto de escolha dividido em blocos).

Tabela 8.6 – Estimativas dos Coeficientes

Atributos	LMPC		LMN	
	Coeficiente	Teste t	Coeficiente	Teste t
Estacionamento	0,9815	6,7938	0,9944	6,8136
Lojas de Conveniência	0,6775	4,7269	0,6779	4,6701
Organização	0,9713	6,7490	0,9740	6,6978
Sistema crédito	1,1096	7,7853	1,1167	7,7648
Níveis serviços	0,9700	6,6988	0,9687	6,6200
Diversificação	0,8009	5,5566	0,7945	5,4537
Desconto	0,6659	4,5962	0,6503	4,4236
$\mathcal{L}(0)$ e $\mathcal{L}(\beta)$	-716,510	-550,984	-716,510	-550,956
LR	331,0511		331,1072	
ρ^2	0,2310		0,2311	
Eficiência = 0,9143				

<i>LMPC – Matriz Covariâncias ($\times 10^{-2}$)</i>	<i>LMN – Matriz Covariâncias ($\times 10^{-2}$)</i>
2,09 1,15 1,12 1,08 1,08 1,14 1,14	2,13 1,15 1,11 1,07 1,08 1,13 1,12
1,15 2,05 1,09 1,14 1,11 1,14 1,12	1,15 2,11 1,10 1,14 1,09 1,12 1,10
1,12 1,09 2,07 1,12 1,17 1,11 1,07	1,11 1,10 2,11 1,11 1,15 1,09 1,07
1,08 1,14 1,12 2,03 1,15 1,10 1,13	1,07 1,14 1,11 2,07 1,14 1,10 1,12
1,08 1,11 1,17 1,15 2,10 1,06 1,11	1,08 1,09 1,15 1,14 2,14 1,06 1,11
1,14 1,14 1,11 1,10 1,06 2,08 1,15	1,13 1,12 1,09 1,10 1,06 2,12 1,14
1,14 1,12 1,07 1,13 1,11 1,15 2,10	1,12 1,10 1,07 1,12 1,11 1,14 2,16

Valor do $\chi^2_{(7)}$: Limite Excelente (5%) : 2,17 ; Limite de aceitação (95%) : 14,1

Teste da Região de Confiança: RC = 0,054

Resultado: => Aceita-se a hipótese de que os coeficientes estimados por LMN estão incluídos no I.C. dos coeficientes estimados por LMPC.

Teste de Comparação: HM = 0,76; MTT = 0,056

Resultado: => Aceita-se a hipótese de igualdade de coeficientes e o ajuste é EXCELENTE.

Descrição dos resultados: As estimativas dos coeficientes (LMPC e LMN) são estatisticamente iguais. Grande parte dessa proximidade é devida ao fato de que o conjunto de escolha formado por arranjos ortogonais foi dividido, utilizando-se das técnicas de BIB; no entanto, somente o Modelo LMPC garante que a propriedade IID é preservada e a probabilidade condicional é levada em consideração.

8.2. DADOS DE PREFERÊNCIA REVELADA

Durante a realização da entrevista, foi solicitado que o entrevistado indicasse, na ordem de preferência, dois supermercados onde normalmente realiza suas compras mensais. De posse desses dados foi possível montar o delineamento constante da tabela 8.7.

Tabela 8.7 – Delineamento Especial com dados PR

Alternativas (Supermercados)	Atributos / Níveis				
	X ₁	X ₂	X ₃	X ₄	X ₅
SM-A1	1	0	0	0	0
SM-A2	0	1	0	0	0
SM-B1	0	0	1	0	0
SM-B2	0	0	0	1	0
SM-A3	0	0	0	0	1

A tabela 8.8 apresenta as estimativas dos parâmetros de delineamento experimental especial, geral e segmentada por renda familiar. Os parâmetros foram estimados através do modelo Logit Multinomial, com estimativas de log-verossimilhança e método de Newton-Raphson. Levando-se em conta que os atributos presentes no experimento são exclusivos, os parâmetros representam diretamente a utilidade que a população amostrada atribui aos respectivos supermercados, U_i^{PR} .

*Tabela 8.8 – Estimativa dos coeficientes – Experimento com dados PR
(Geral e por segmentação de Renda)*

Super Mercados	Coeficientes = Utilidades dos Supermercados				
	Geral	Classe D	Classe C	Classe B	Classe A
SM-A1	2,280 (5,70)	1,688 (1,74)	1,966 (2,66)	2,943 (3,78)	2,759 (2,74)
SM-A2	2,403 (6,05)	1,518 (1,54)	2,731 (3,82)	2,107 (2,64)	3,356 (3,38)
SM-B1	3,072 (7,82)	4,679 (4,95)	3,581 (5,02)	2,581 (3,29)	2,016 (1,96)
SM-B2	1,708 (4,13)	2,924 (3,20)	1,478 (1,92)	0,883 (0,98)	0,733 (0,62)
SM-A3	2,314 (5,81)	1,677 (1,72)	1,705 (2,25)	3,068 (3,98)	2,568 (2,55)
N. amostra	175 (350)	39 (78)	56 (112)	50 (100)	30 (60)
LR	188,38 **	112,58 **	102,45 **	75,12 **	52,14 **
ρ^2	0,1583	0,4244	0,2690	0,2286	0,2552

** - Significativo com 99% de probabilidade;

Notação: Classe D : renda familiar até R\$ 600;
 Classe C : renda familiar de R\$ 601 a 1200;
 Classe B : renda familiar de R\$ 1201 a 2400;
 Classe A : renda familiar acima de R\$ 2400.

A tabela 8.9 apresenta os dados proporcionais dos coeficientes estimados e que serão utilizados na comparação com os coeficientes proporcionais dos dados PD.

Tabela 8.9 – Coeficientes proporcionais estimados : \tilde{U}_i^{PR} .

	Geral	Classe D	Classe C	Classe B	Classe A
SM-A1	19,36	13,52	17,15	25,41	24,13
SM-A2	20,40	12,16	23,83	18,19	29,35
SM-B1	26,08	37,47	31,25	22,28	17,63
SM-B2	14,50	23,42	12,90	7,62	6,41
SM-A3	19,64	13,43	14,88	26,50	22,46

8.3. VALIDAÇÃO EXTERNA DOS DADOS : Integração PR x PD

8.3.1. Avaliação real dos Atributos

Com a finalidade de testar a validação externa dos resultados obtidos através da pesquisa de preferência declarada e determinar os coeficientes de atratividade (CAS), foram avaliados os atributos reais referentes aos delineamentos de *preço*, *qualidade* e *conveniências*, em 4 lojas de supermercados do tipo A e 4 lojas do tipo B, que operam em Guarapuava, cujos resultados constam da tabela 6.3. Os valores das avaliações, quando aplicados nas respectivas FU definidas no item 8.1.2, geram os valores das utilidades, (U_i^{PD}), de cada supermercado do ponto de vista dos clientes de forma geral e segmentados por renda.

8.3.2. Comparação entre as Utilidades: \tilde{U}_i^{PD} vs \tilde{U}_i^{PR}

a) Utilidade geral dos supermercados

A tabela 8.10 e a figura 8.1 apresentam os resultados das aplicações dos valores dos atributos referentes aos supermercados na FU (geral), obtendo-se as utilidades de cada supermercado e possibilitando-se comparar, em unidades comuns, com as utilidades advindas de dados de preferência revelada.

Tabela 8.10 – Utilidade dos Supermercados (Visão Geral dos Clientes)

Geral	PREÇO		QUALIDADE		CONVENIÊNCIA		\tilde{U}_i^{PD}	\tilde{U}_i^{PR}
	2,44		4,21		1,88			
	Coef.	%	Coef.	%	Coef.	%		
SM-A1	0,895	8,36	8,457	26,60	5,510	29,34	21,68	19,36
SM-A2	1,429	13,34	8,382	26,36	4,797	25,54	22,24	20,40
SM-A3	0,667	6,23	7,113	22,37	4,801	25,56	18,19	19,64
SM-B1	4,343	40,55	5,422	17,05	1,943	10,35	22,69	26,08
SM-B2	3,376	31,52	2,419	7,61	1,732	9,22	15,20	14,50
Teste de χ^2 entre os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} =							1,05 (*)	

(*) Como o valor χ^2 calculado é menor que o valor tabelado ($\chi^2_{(4, 95\%)} = 9,45$), então aceita-se H_0 : os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} são estatisticamente iguais.

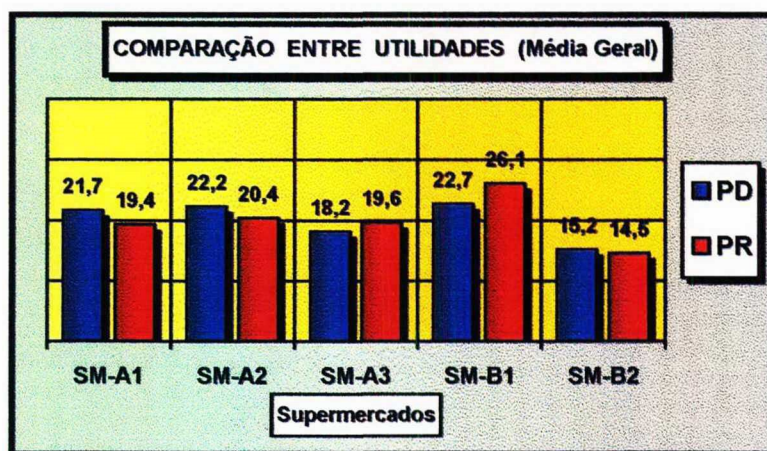


Figura 8.1 – Comparação entre as Utilidades: \tilde{U}_i^{PD} vs \tilde{U}_i^{PR} (Visão geral dos Clientes)

Descrição dos resultados: Observa-se na figura 8.1 um ajuste muito bom entre os dados obtidos na pesquisa de preferência declarada e dados obtidos por preferência revelada. Conclui-se que a Função Utilidade definida pela composição de funções representa muito bem o comportamento geral dos consumidores.

b) *Utilidade dos supermercados: segmentação por Renda Familiar (classe D - renda inferior a R\$ 600);*

A tabela 8.11 e a figura 8.2 apresentam os resultados das aplicações dos valores dos atributos referentes aos supermercados na FU (Classe D), obtendo-se as utilidades de cada supermercado e possibilitando comparar, em unidades comuns, com as utilidades advindas de dados de preferência revelada.

Tabela 8.11 – Função Utilidade dos Supermercados (Renda familiar até R\$ 600)

Mercados	PREÇO		QUALIDADE		CONVENIÊNCIA		\tilde{U}_i^{PD}	\tilde{U}_i^{PR}
	x 2,04		X 4,98		x 2,14			
SM-A1	1,0275	8,75	7,8411	26,53	5,6706	30,17	21,65	13,52
SM-A2	1,5732	13,40	7,7582	26,25	4,9325	26,24	22,05	12,16
SMA-3	0,3722	3,17	6,6147	22,38	4,9731	26,46	17,14	13,43
SM-B1	5,1674	44,01	5,0465	17,07	1,6985	9,04	23,83	37,47
SM-B2	3,6018	30,67	2,2977	7,77	1,5219	8,10	15,32	23,42
Teste de χ^2 entre os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} =							20,5 (*)	

(*) Como o valor χ^2 calculado é maior que o valor tabelado ($\chi^2_{(4; 95\%)} = 9,45$), então rejeita-se H_0 : os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} são estatisticamente diferentes.

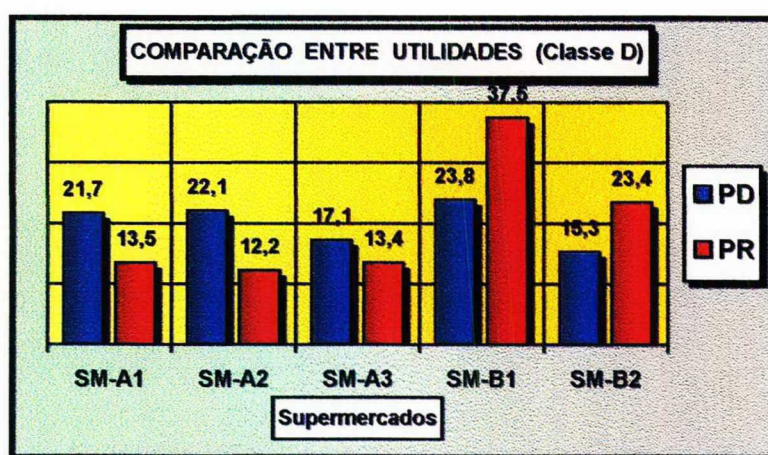


Figura 8.2 – Comparação entre as Utilidades: \tilde{U}_i^{PD} vs \tilde{U}_i^{PR}
(Visão dos Clientes com renda familiar até R\$ 600)

Descrição dos resultados: A figura 8.2 demonstra um ajuste razoável entre os dados obtidos na pesquisa de preferência declarada e dados obtidos por preferência revelada, quando segmentado por renda (Classe D – renda familiar inferior a R\$ 600). Existe uma vantagem real de utilização dos supermercados do tipo B em relação ao tipo A. Percebe-se, ainda, que a amostra da classe A tem uma inspiração por uma situação melhorada, mas na prática, as ações passam pelas questões econômicas e realizam suas compras nos supermercados do tipo B.

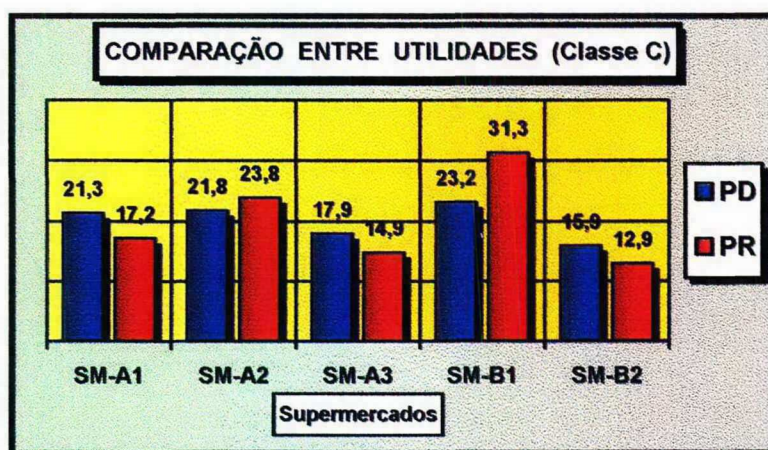
- c) Utilidade dos supermercados: segmentação por Renda Familiar (classe C - renda de R\$ 601 a 1200);

A tabela 8.12 e a figura 8.3 apresentam os resultados das aplicações dos valores dos atributos referentes aos supermercados na FU (Classe C), obtendo-se as utilidades de cada supermercado e possibilitando comparar, em unidades comuns, com as utilidades advindas de dados de preferência revelada.

*Tabela 8.12 – Função Utilidade dos Supermercados
(Visão dos Clientes com renda familiar de R\$ 601 a 1200)*

Mercados	PREÇO		QUALIDADE		CONVENIÊNCIA		\tilde{U}_i^{PD}	\tilde{U}_i^{PR}
	x 2,36		x 3,62		x 2,18			
SM-A1	0,6267	6,38	7,6841	26,50	5,7698	29,45	21,28	17,15
SM-A2	1,2475	12,69	7,5666	26,10	4,8825	24,92	21,78	23,83
SMA-3	0,4383	4,46	6,4785	22,35	5,0288	25,67	17,89	14,88
SM-B1	4,245	43,19	4,9417	17,04	2,0978	10,71	23,16	31,25
SM-B2	3,2709	33,28	2,3216	8,01	1,8126	9,25	15,88	12,90
Teste de χ^2 entre os valores de \tilde{U}_i^{PD} e $\tilde{U}_i^{PR} =$							4,86 (*)	

(*) Como o valor χ^2 calculado é menor que o valor tabelado ($\chi^2_{(4; 95\%)} = 9,45$), então aceita-se H_0 : os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} são estatisticamente iguais.



*Figura 8.3 – Comparação entre as Utilidades: \tilde{U}_i^{PD} vs \tilde{U}_i^{PR}
(Visão dos Clientes com renda familiar de R\$ 601 a 1200)*

Descrição dos resultados: A figura 8.3 demonstra um bom ajuste entre os dados obtidos na pesquisa de preferência declarada e dados obtidos por preferência revelada, quando segmentado por renda (Classe C – renda familiar de R\$ 601 a 1200).

d) Utilidade dos supermercados: segmentação por Renda Familiar (classe B - renda de R\$ 1201 a 2400);

A tabela 8.13 e a figura 8.4 apresentam os resultados das aplicações dos valores dos atributos referentes aos supermercados na FU (Classe B), obtendo-se as utilidades de cada supermercado e possibilitando comparar, em unidades comuns, com as utilidades advindas de dados de preferência revelada.

*Tabela 8.13 – Função Utilidade dos Supermercados
(Visão dos Clientes com renda familiar de R\$ 1201 a 2400)*

Mercados	PREÇO		QUALIDADE		CONVENIÊNCIA		\tilde{U}_i^{PD}	\tilde{U}_i^{PR}
	x 2,71		x 4,49		x 1,38			
SM-A1	0,4388	4,01	9,407	26,75	5,3403	30,38	19,87	25,41
SM-A2	1,6473	15,04	9,3356	26,55	4,2523	24,19	22,42	18,19
SM-A3	1,0501	9,59	7,8492	22,32	4,3586	24,80	18,54	26,50
SM-B1	4,3124	39,38	5,9678	16,97	1,9399	11,04	23,38	22,28
SM-B2	3,5028	31,98	2,6074	7,41	1,6861	9,59	15,79	7,62
Teste de χ^2 entre os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} =							10,06 (*)	

(*) Como o valor χ^2 calculado é maior que o valor tabelado ($\chi^2_{(4; 95\%)} = 9,45$), então rejeita-se H_0 : os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} são estatisticamente diferentes.

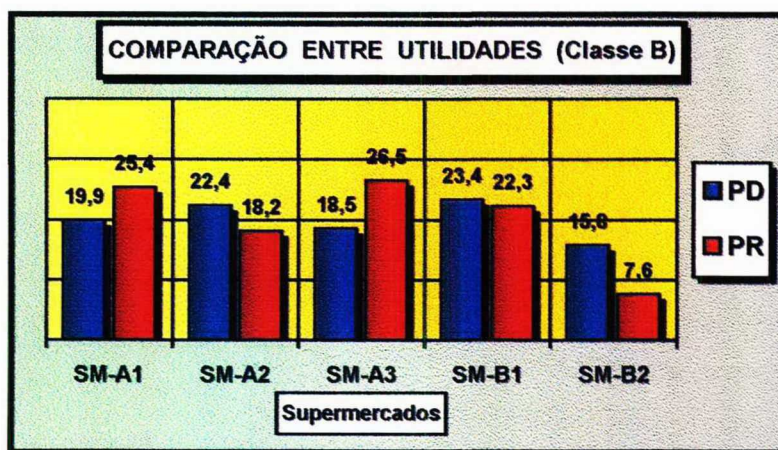


Figura 8.4 – Comparação entre as Utilidades: \tilde{U}_i^{PD} vs \tilde{U}_i^{PR}

Descrição dos resultados: A figura 8.4 demonstra um ajuste razoável entre os dados obtidos na pesquisa de preferência declarada e dados obtidos por preferência revelada, quando segmentado por renda (Classe B). Percebe-se que nessa faixa de renda existe uma vantagem de preferência para as compras nos supermercados do tipo A, potencializado nas estimativas com os dados de preferência revelada.

- e) Utilidade dos supermercados: segmentação por Renda Familiar (classe A - renda acima de R\$ 2400);

A tabela 8.14 e a figura 8.5 apresentam os resultados das aplicações dos valores dos atributos referentes aos supermercados na FU (Classe A), obtendo-se as utilidades de cada supermercado e possibilitando comparar, em unidades comuns, com as utilidades advindas de dados de preferência revelada.

*Tabela 8.14 – Função Utilidade dos Supermercados
(Visão dos Clientes com renda familiar acima de R\$ 2400)*

Mercados	PREÇO		QUALIDADE		CONVENIÊNCIA		\tilde{U}_i^{PD}	\tilde{U}_i^{PR}
	x 2,64		X 4,83		x 1,25			
SM-A1	-0,3523	0,00	11,600	26,77	5,4058	30,22	19,15	24,13
SM-A2	0,3428	14,73	11,652	26,89	4,1704	23,32	22,69	29,35
SM-A3	0,6006	20,19	9,776	22,56	4,4418	24,83	22,17	22,46
SM-B1	0,9618	27,85	7,456	17,21	1,9765	11,05	19,56	17,63
SM-B2	1,4047	37,23	2,847	6,57	1,8916	10,58	16,43	6,41
Teste de χ^2 entre os valores de \tilde{U}_i^{PD} e $\tilde{U}_i^{PR} =$							9,50 (*)	

(*) Como o valor χ^2 calculado é menor que o valor tabelado ($\chi^2_{(4; 95\%)} = 9,45$), então, rejeita-se H_0 : os valores de \tilde{U}_i^{PD} e \tilde{U}_i^{PR} são estatisticamente diferentes.

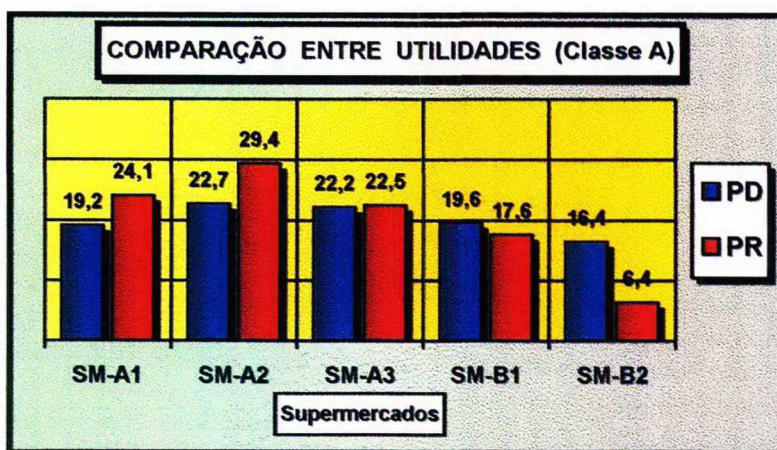


Figura 8.5 – Comparação entre as Utilidades: \tilde{U}_i^{PD} vs \tilde{U}_i^{PR}

Descrição dos resultados: A figura 8.5 demonstra um bom ajuste entre os dados obtidos na pesquisa de preferência declarada e dados obtidos por preferência revelada, quando segmentado por renda (Classe A). Percebe-se que nessa faixa de renda existe uma vantagem de preferência para as compras nos supermercados do tipo A.

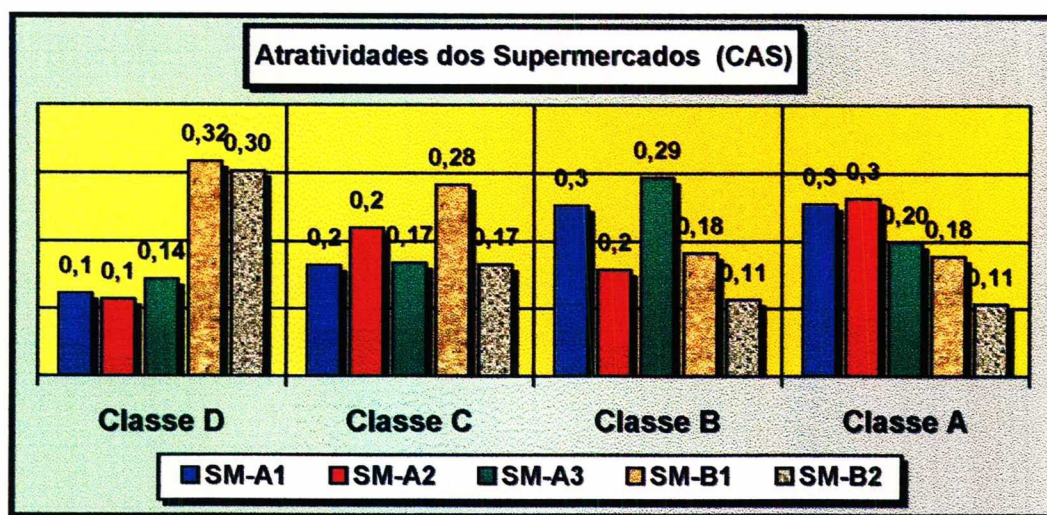
8.3.3. Coeficiente de Atratividade dos Supermercados (CAS)

Quando são substituídos, na FU, os valores avaliados para os atributos, cada supermercado define sua Utilidade dada por PD. Os CAS indicam uma relação de atratividade entre os supermercados além das utilidades definidas por PD, ou seja, é um coeficiente de utilidade não explicada pela FU advinda da PD e sim definida por outros atributos.

A tabela 6.15 apresenta os valores dos Coeficientes de Atratividade dos Supermercados, segmentado por classes de renda familiar e calculados conforme descrito no item 6.2.5.

*Tabela 8.15 – Coeficiente de Atratividade dos Supermercados
(Segmentado por classes de renda familiar)*

	Classe D	Classe C	Classe B	Classe A
SM-A1	0,1226	0,1642	0,2523	0,2543
SM-A2	0,1140	0,2190	0,1581	0,2628
SM-A3	0,1438	0,1685	0,2932	0,1986
SM-B1	0,3165	0,2828	0,1822	0,1776
SMB2	0,3028	0,1652	0,1138	0,1065



*Figura 8.6 – Coeficientes de Atratividade dos Supermercados
(Segmentado por classes de rendas)*

Descrição dos resultados: Percebe-se, de acordo com a figura 8.6, que existe uma vantagem de atratividade dos supermercados do tipo B em relação aos do tipo A, na classe D. Essa vantagem vai se modificando ao passar pelas classes e na Classe A há uma inversão de atratividade, com vantagem para os supermercados do tipo A.

8.4. DISCUSSÃO DOS RESULTADOS DA APLICAÇÃO

Com o término dos trabalhos de análise e interpretação dos dados da pesquisa de campo que envolveram dados de preferência declarada e revelada, podem-se destacar os seguintes aspectos:

8.4.1. Comportamento do modelo adotado para a análise

Em primeiro lugar, cabe ressaltar que a conceituação, organização e a estruturação da pesquisa facilitou a construção dos delineamentos experimentais e a confecção do questionário que ficou simples, claro e objetivo, evitando perda de qualidade das informações por desmotivação ou fadiga (Anexo 7).

Para atingir os objetivos propostos, utilizaram-se os modelos: Logit Hierárquico, Logit Multinomial e Logit Multinomial com probabilidade condicional, desenvolvido nesta tese, e a aplicação da função Logit para determinar os coeficientes da análise de regressão múltipla com dados de preferência revelada. As estimativas dos parâmetros e testes de hipóteses foram calculadas utilizando-se o software denominado LMPC desenvolvido em ambiente Pascal (Delphi) pelo autor da tese.

As análises com dados de preferência declarada apresentaram um comportamento assintótico e convergente, em que o ponto de máximo da função foi obtido com até 5 iterações, através do processo de Newton-Raphson com busca unidirecional por seção áurea e apresentando variabilidades compatíveis com os dados. Dessa forma pode-se concluir que o comportamento do modelo adotado esteve dentro do esperado e atende aos propósitos da pesquisa.

A conclusão admitida no parágrafo anterior tem seu valor corroborado quando esses resultados obtidos através dos métodos de preferência declarada são comparados com os resultados obtidos pelo método de preferência revelada, conforme se observa no item 8.3.3.

8.4.2. Amostragem e análise dos dados

Foram coletadas 175 amostras válidas, cujos dados descritivos estão tabelados no item 6.1.

Quanto à amostragem, os dados de preferência declarada foram filtrados mediante os seguintes procedimentos:

- (a) *Tratamento dos dados discrepantes (TDD)* – identifica o comportamento assimétrico das informações e possíveis entrevistas discrepantes, fora do padrão normal e com sugestão para eliminação conforme ocorreu no delineamento experimental *QUALIDADE*; nos demais casos o comportamento foi normal e simétrico;
- (b) *Dimensionamento da amostra* – procedimento para identificar se a quantidade amostrada possui dados suficientes, em função de sua variabilidade, para estimar coeficientes assintoticamente convergentes. Em todas as análises o dimensionamento mostrou-se satisfatório;
- (c) *Transformação dos dados* – procedimento que identifica a existência de heterocedasticidade regular intra-alternativa e indica, conforme ocorreu no delineamento *QUALIDADE*, a transformação a ser utilizada. Após a utilização da transformação, os dados mostraram-se homocedásticos, ou seja, a heterocedasticidade regular constatada, que ocasiona viés nos resultados, foi corrigida. Nos demais delineamentos não foi constatada a heterocedasticidade regular.

Quanto aos testes das estatísticas:

- (a) *Razão de verossimilhança (LR)* – todos se apresentaram significativos com 99% de confiabilidade;
- (b) *Coefficiente de correlação* (para função de verossimilhança) ρ^2 – tiveram seus valores entre o mínimo de 0,23 e máximo de 0,48, para os dados de PD e de 0,17 a 0,42, respectivamente, para os dados de PR, indicando uma correlação satisfatória, tendo em vista que 0,40 seria o ótimo.

Assim, com as condições amostrais, definidas como satisfatórias, foi possível construir as Funções Utilidades, descritas no item 8.1.2. Dos valores das avaliações dos atributos relativos aos supermercados em Guarapuava, aplicados na FU, obteve-se o valor de utilidade dos respectivos supermercados, de forma geral e segmentado por renda, de acordo com a população amostrada. Essa informação foi confrontada com os resultados dos dados de preferência revelada, segundo o paradigma descrito na figura 6.4, como forma de validação da análise e, em um segundo momento, para determinar os Coeficientes de Atratividade dos Supermercados. Conforme as observações descritas após o figura 8.1, existe, de forma geral, um ajuste muito bom entre os dados obtidos na pesquisa de preferência declarada e dados obtidos por preferência revelada. Conclui-se que a Função Utilidade definida pela composição de funções representa muito bem o comportamento dos consumidores.

8.4.3. Resultados obtidos a respeito das utilidades dos supermercados

Os resultados obtidos podem ser analisados seguindo duas óticas:

a) Comportamento da população amostrada.

Dentre os atributos sociais levantados, constatou-se que a segmentação por classes de renda é a que melhor agrupa as amostras, segundo as variabilidades, formando segmentos mais homogêneos para o tipo de análise objetivada.

Nas observações descritas a respeito dos dados comparativos entre U_i^{PD} e U_i^{PR} , de forma geral e segmentados por classes de renda familiar, pode ser constatado o seguinte: *Existe um comportamento diferenciado na população amostrada, quando a amostra é segmentada por classes de renda.* A tabela 8.16 e a figura 8.7 descrevem de forma gráfica a afirmativa citada anteriormente.

*Tabela 8.16 – Dados percentuais das utilidade:
por tipo de mercado e por segmentação de renda*

	Classe D		Classe C		Classe B		Classe A		Geral	
	Tipo A	Tipo B	Tipo A	Tipo B	Tipo A	Tipo B	Tipo A	Tipo B	Tipo A	Tipo B
U_i^{PD}	50,9	49,1	51,0	49,0	50,9	49,1	54,2	45,8	52,2	47,8
U_i^{PR}	30,0	70,0	45,8	54,2	61,0	39,0	67,8	32,2	49,4	50,6

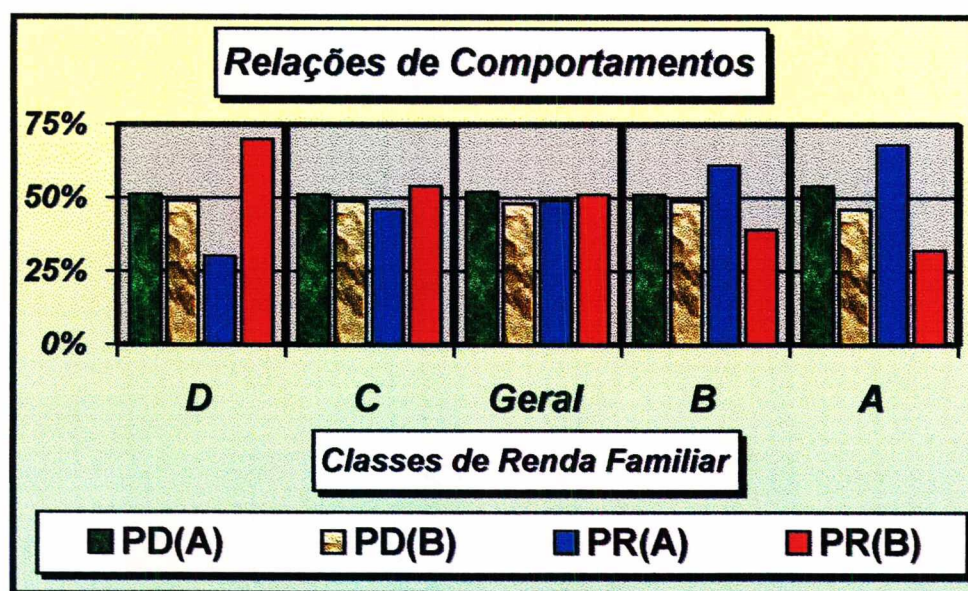


Figura 8.7 – Relações de Comportamentos

Notação: $PD(A)$ – Média das U_i^{PD} , dos supermercados classificados no Tipo A;
 $PD(B)$ – Média das U_i^{PD} , dos supermercados classificados no Tipo B;
 $PR(A)$ – Média das U_i^{PR} , dos supermercados classificados no Tipo A;
 $PR(B)$ – Média das U_i^{PR} , dos supermercados classificados no Tipo B;

Uma das conclusões que se pode extrair da figura 8.7 a respeito do comportamento da população amostrada é que as pessoas, quando inquiridas a respeito de suas opiniões, a serem expressas, através da ordenação dos cartões (Técnica de Preferência Declarada) revelam um comportamento quase isento das suas condições sociais, ou seja, um comportamento baseado em conceitos adquiridos pelo convívio em sociedade, talvez influenciado pelos meios de comunicações e que apresentam diferenças em relação ao seu comportamento real. Isso explica o comportamento aproximadamente uniforme, entre as classes sociais, no que se refere à valorização dos atributos e definição da utilidade dos supermercados, conforme pode ser observado, no gráfico, na relação uniforme entre as áreas definidas por $PD(A)$ e $PD(B)$, no decorrer das classes de renda familiar. Mas, quando realizam as ações em função de sua condição social, buscam otimizar os recursos, externando comportamentos diferenciados conforme a classe de renda.

Por exemplo, observa-se, no gráfico, que indivíduos da classe D, com renda até R\$ 600, e da classe A, com renda superior a R\$ 2400, apresentam comportamentos semelhantes quando analisados com dados de PD $\{PD(A) \text{ e } PD(B)\}$ e comportamentos diferentes quando analisados com dados de PR $\{PR(A) \text{ e } PR(B)\}$. Os indivíduos da classe D preferem comprar nos Supermercados do tipo B e os da classe A preferem fazer suas compras nos supermercados do tipo A. As classes intermediária vão alternando de comportamento e na análise geral os comportamentos são semelhantes.

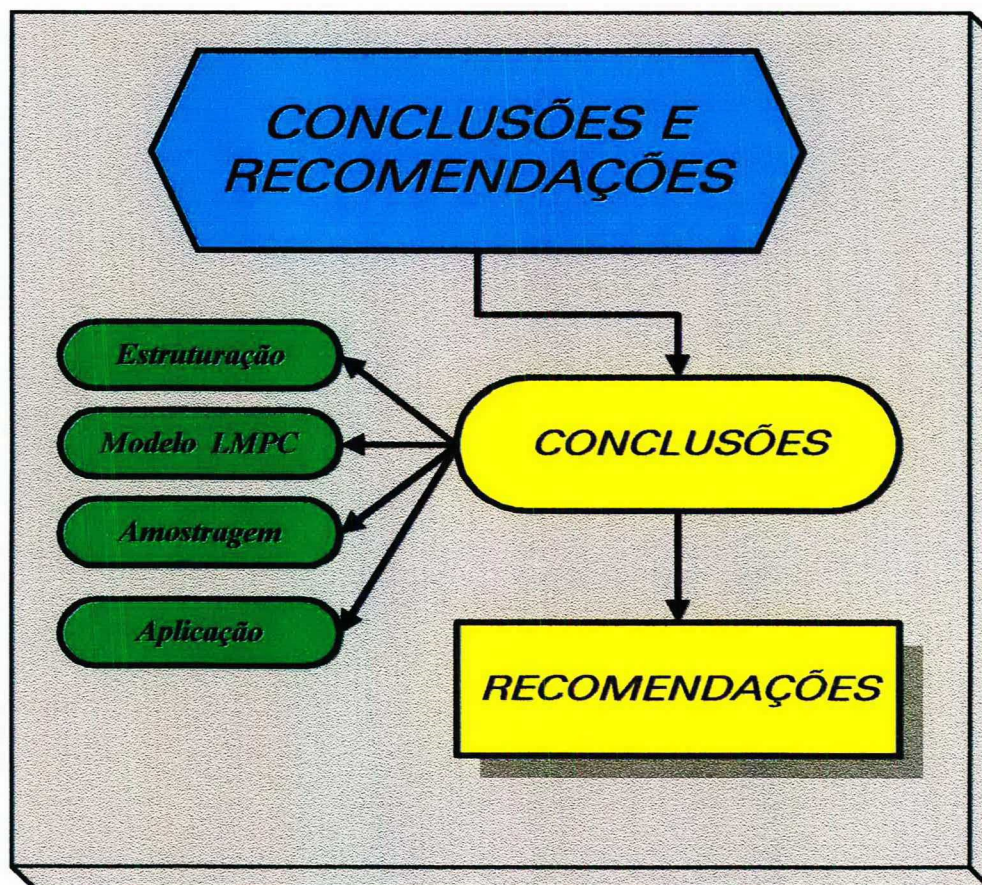
b) Importância do conhecimento dos resultados

A importância que o conhecimento dos resultados: da FU, dos dados comparativos, do comportamento da população e dados de benchmarking, traz para uma administração voltada para o cliente e que tenha compromisso com a imagem da empresa é a grande vantagem competitiva, ou seja, é o caminho da diferenciação que conduz à excelência .

Em uma administração com qualidade total, não se trata apenas de aumentar a qualidade e a produtividade, de reduzir custos e de racionalizar procedimentos de produção e transporte, mas o fator crítico está na percepção de *valor*. Valor é muito mais do que preço. É um conjunto de benefícios que o produto, com vantagens e atrativos, pode oferecer aos clientes. Mais do que o produto em si e sua qualidade, o valor é fruto de percepção, informação e experimentação, nem sempre combinadas. Assim, de posse do conhecimento da relação entre os atributos que expressam a opinião dos clientes, e conhecedor de suas potencialidades é possível estabelecer um plano ótimo de investimentos que venham revestir sua empresa de acordo com as aspirações dos clientes em busca da excelência.

PARTE III – CONCLUSÃO

Capítulo 9 – CONCLUSÕES E RECOMENDAÇÕES



9. CONCLUSÕES E RECOMENDAÇÕES

9.1. CONCLUSÕES

As principais dificuldades enfrentadas pelos pesquisadores, em experimentos com preferência declarada, estão relacionadas com a estruturação do problema e essa estrutura está vinculada à definição de uma função matemática que deverá representar o problema real. A configuração de um problema real geralmente envolve um número muito grande de variáveis, caracterizando a complexidade do modelo, o que impõe uma reflexão sobre a estruturação do problema.

Tendo em vista que os métodos de pesquisa de preferência declarada utilizam respostas individuais a respeito da preferência em um conjunto de opções, de forma a estimar funções utilidade, deve-se estabelecer um equilíbrio entre a complexidade do experimento e a capacidade humana de interpretação. O presente trabalho desenvolve dois procedimentos complementares com o objetivo de estruturar problemas complexos, que envolvem muitos atributos e níveis, de forma compatível com a capacidade de interpretação humana, segundo o número de Miller:

- (a) metodologia da divisão do problema complexo real em vários problemas menores e mais homogêneos, numa estrutura de árvore, conforme a figura 5.1, baseado na filosofia “dividir para conquistar”. Nesse procedimento há necessidade de se estabelecer uma forma de relacionamento entre os ramos ou áreas de interesse para a composição em uma função geral dos resultados. Há uma diferença básica em relação às estruturas referentes ao modelo logit hierárquico, uma vez que os ramos da árvore constituem-se em experimentos independentes e não em níveis hierárquicos de um único experimento;

(b) processos de otimização na construção do conjunto das alternativas com a incorporação do arranjos ortogonais de Taguchi em pesquisa de preferência declarada, além de estruturar a utilização dos arranjos fatoriais fracionários e técnicas de confundimento para soluções similares. Em complementação, foi organizado um esquema para utilização da divisão do conjunto de alternativas em conjuntos menores, denominados conjuntos de escolha, compatíveis com o tipo de pesquisa a ser efetivada, utilizando-se das técnicas de blocos incompletos: BIB, BIPB e BINB.

Como a utilização do conjunto de escolha, formado pela divisão do conjunto de alternativas, impõe a necessidade de o modelo matemático levar em consideração a probabilidade condicional para atender o pressuposto da propriedade IID, desenvolveu-se o modelo Logit Multinomial com Probabilidade Condicional (LMPC). No modelo LMPC a probabilidade condicional advém do processo sistemático de apresentação dos conjuntos de escolhas aos entrevistados. Neste caso, há a necessidade de preservar a igualdade de representação de cada conjunto de escolha na constituição dos dados. Este procedimento garante que o processo de ordenação nos conjuntos de escolha estimam, sem viés, a ordenação de todas as alternativas e, por conseguinte, a estimativa dos parâmetros. A seção 3.4 apresenta o desenvolvimento do modelo, e os aspectos computacionais constam do anexo 6.

Foram desenvolvidos, também de forma inédita, outros 3 procedimentos a serem adotados para a validação da amostragem dos dados: dimensionamento da amostra (DA); tratamento de dados discrepantes (TDD) e transformação dos dados quando existe heterocedasticidade regular intra-alternativa. (capítulo 4).

A validação da estruturação, do modelo e dos testes de amostragem ocorreram mediante uma aplicação prática em pesquisa na área de marketing (capítulo 6, 7, e 8). Dos resultados obtidos, cabe ressaltar alguns aspectos importantes:

a) Quando à metodologia adotada para a pesquisa de preferência declarada:

A tabela 9.1 resume os problemas (objetivos) e os procedimentos ou soluções adotadas na aplicação:

Tabela 9.1 – Problemas e soluções

Problemas	Procedimentos -> solução		
1 - Montar um experimento com 15 atributos: $3^4 \times 2^{11}$ com 165888 alternativas.	O experimento foi estruturado em árvore, conforme a figura 5.1, onde cada ramo constituiu-se em uma variável latente: Preço = 3^4 ; Qualidade = 2^4 ; Conveniência = 2^7 . Foi montado um experimento para estabelecer a forma de relacionamento entre as variáveis latentes.		
	PREÇO $3^4 = 81$ alternativas	QUALIDADE $2^4 = 16$ alternativas	CONVENIÊNCIA $2^7 = 128$ alternativas
2 – Otimizar o conjunto de alternativas.	Foi montado um experimento hierárquico em dois níveis, o 2º nível ficou com $3^3 = 27$ alternativas.	Utilizou-se do arranjo ortogonal de Taguchi L8 = 8 alternativas. Excluiu-se duas alternativas por serem totalmente dominada e dominante, assim o conjunto de escolha ficou com 6 alternativas.	Utilizou-se do arranjo ortogonal de Taguchi L8 = 8 alternativas.
3 – Otimizar o conjunto de escolha.	O conjunto de escolha da 2ª etapa ficou com 3 alternativas, formadas pela técnica de blocos de cadeia generalizada (BCG).		Resultou 7 conjunto de escolha com 4 alternativas cada. Foram formados pela divisão do conjunto das alternativas, utilizando-se das técnicas de blocos incompletos balanceados (BIB).
4 – Amostragem: • Dados discrepantes: • Dimensionamento: • Transformação:	Resultados: → -*- → 137 entrevistas; → sem transformação.	Resultados: → 21 discrepantes; → 130 entrevistas; → logarítmica.	Resultados: → 0 discrepantes; → 135 entrevistas; → sem transformação.
5 – Estimativas dos parâmetros	Logit Multinomial com Probabilidade condicional (LMPC)	Logit Multinomial (LMN).	Logit Multinomial com Probabilidade Condicional (LMPC)
6 – Testes: • LR • ρ^2 • ρ^2_{Aj}	Resultados: 251,6 ** 0,2893 (x) 0,2801 (x)	Resultados: 178,3 ** 0,2814 (x) 0,2688 (x)	Resultados: 331,0 ** 0,2310 (x) 0,2212 (x)

** Significativo com 95% de confiabilidade (Todos coeficientes são diferentes de zero);

(x) Os valores para os coeficientes de determinação (ρ^2 e ρ^2_{Aj}) entre 0,2 e 0,4 são considerados significativos.

b) Quanto aos resultados da aplicação:

Os resultados da aplicação foram descritos no item 8.4 em que se concluiu que a Função Utilidade definida pela composição das funções, que utilizou-se dos procedimentos descritos nesta tese, representa muito bem o comportamento geral dos consumidores (item 8.3 e figura 8.1), embora esses comportamentos sejam diferenciados quando segmentados por classes de rendas.

Os resultados obtidos através das técnicas de preferência declarada, comparados com os resultados da pesquisa de preferência revelada, expressaram uma aspiração intrínseca da população, enraizada de forma independente de suas condições sociais. No entanto, as ações realizadas em função da condição social buscam otimizar os recursos, externando comportamentos diferenciados conforme as classes de rendas (figura 8.7).

c) Quanto à utilidade do modelo LMPC:

A utilidade do modelo LMPC tem seu mérito evidenciado pela demonstração da fragilidade do modelo LMN em tratar com experimentos cujos conjuntos de escolhas são oriundos da divisão de um conjunto maior de alternativas, em que se impõe a probabilidade condicional para o atendimento do pressuposto da propriedade IID. Nas duas oportunidades em que foram comparados (item 8.1.3), obtiveram-se menores variabilidades para o modelo LMPC em relação ao modelo LMN. BEN-AKIVA e BOCCARA (1995) também obtiveram conclusões semelhantes. Dessa forma o modelo LMPC é recomendado quando se leva em conta a probabilidade condicional no modelo de escolha discreta.

d) Quanto aos processos de estimação:

Foi observada a necessidade da inclusão de uma nova regra de parada a ser implementada no processo de maximização de Newton-Raphson. Essa nova regra de parada leva em conta a maximização do valor da estatística t de Student nos parâmetros e visa não permitir erros de estimação advindos de possíveis combinações lineares, o que resultaria em ponto de quase-cumeieira da FU (Anexo 6). BEN-AKIVA e

BOCCARA (1995) já alertaram a respeito de possíveis problemas na concavidade do ponto de máximo para a função de log-verossimilhança com probabilidade condicional.

e) Quanto aos testes de verificação da amostra

A aplicação desses procedimentos mostrou-se válida, conforme consta das discussões dos resultados apresentadas no capítulo 7, uma vez que se passou a conhecer melhor os dados amostrais, transmitindo maior confiabilidade aos resultados.

- *Dimensionamento da Amostra:*

A utilização do procedimento para o dimensionamento da amostra, quando aplicado ao total amostrado, tem uma função importante na indicação da fidedignidade das estimativas dos parâmetros em relação à população amostrada.

- *Tratamento dos dados discrepantes:*

Constitui-se em um procedimento eficiente para a identificação das entrevistas discrepantes em relação ao comportamento geral da população amostrada. A identificação do percentil discrepante é feita em função do Coeficiente de assimetria de Pearson, conforme descrito na seção 4.1. Com a exclusão das entrevista discrepante, os resultados apresentam uma melhoria de performance em relação ao testes LR e ρ^2 .

- *Transformação dos dados:*

Constata-se que o uso da transformação logarítmica pode ser indicada quando existe heterocedasticidade nos dados obtidos pela escolha explodida das alternativas. A heterocedasticidade pode ser consequência de dois fatores: (1) falta a homogeneidade no valor da utilidade atribuída em cada escolha pelo entrevistado, e (2) por questões inerentes à própria distribuição teórica dos erros (Distribuição de Gumbel), identificado no desenvolvimento teórico de Bartlett. A violação da suposição de homocedasticidade pode provocar estimativa do erro com viés, o que vem causar distorções nos testes de validação (teste t, intervalos de confiança e teste de comparação de médias) e nas estimativas dos parâmetros. A transformação produz estimativas do erro, em geral,

maior que nos dados sem transformação, impondo maior rigor no teste t e também produz pequenas variações nos valores das estimativas dos parâmetros β_i . Uma das consequências da não observação da suposição de homogeneidade das variâncias pode ser a de cometer Erro do Tipo I, rejeitar a hipótese de nulidade quando na verdade deveria aceitar.

9.2. RECOMENDAÇÕES

Estudos complementares referentes ao processo de maximização da função de log-verossimilhança para os modelos logit multinomial com probabilidade condicional e logit hierárquico;

Montagem de delineamentos integrando os experimentos hierárquicos com o modelo Logit Multinomial com probabilidade condicional;

Aplicações da estruturação em árvore com incorporação dos arranjos ortogonais de Taguchi e blocos incompletos em pesquisas de preferência declarada nas áreas de transportes ou logística;

Estudos complementares referentes às implicações da transformação nos níveis dos atributos, em pesquisa de preferência declarada, quando eles não se apresentem de forma linear;

Otimização dos investimentos financeiros a serem aplicados de acordo com os resultados de pesquisas de preferência declarada com o objetivo de maximizar o retorno da utilidade.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 ACKERMANN, F.; EDEN, C.; CROPPER, S. Getting started with cognitive mapping. Artigo fornecido com o **Software COPE**, 1995.
- 2 ADDELMAN, S. Orthogonal main-effect plans for asymmetrical factorial experiments. **Technometrics**, v. 4, n. 1 p. 21-58, 1962.
- 3 AMEMYA, t. Qualitative response models. **Ann. Econ. Social Measurement**, n. 4, p. 363-372, 1975.
- 4 ANDERSEN, P. B.; MOELLER, J.; SHELDON, R. J. Marketing DSB rail services using a stated preference approach. **PTRC Summer Annual Meeting**, Bridgeton, n. 14, p. 263-270, July 1986.
- 5 ANDERSON, N. H. Functional measurement and psychological judgment. **Psychological Review**, v. 77, n. 3, p. 153-70, 1970.
- 6 ARTES, Rinaldo. **Análise de preferência "Conjoint analysis"**. São Paulo, 1991. Dissertação de mestrado em Estatística – Instituto de Matemática e Estatística, Universidade de São Paulo.
- 7 BANA E COSTA, C. A. **Struturation, construction et exploitation d'un modèle multicritère d'aide à la décision**. Lisboa, 1992. Tese de Doutorado. Universidade Técnica de Lisboa.
- 8 BANZATTO, David A.; KRONKA, Sérgio do N. **Experimentação agrícola**. Jaboticabal : FUNEP, 1989.
- 9 BARTLETT, Maurice S. The use of transformations. **Biometrics**, v. 3, p. 39-57, 1947.
- 10 BASTOS, Lia Caetano. **Planejamento da rede escolar : uma abordagem utilizando preferência declarada**. Florianópolis, 1994. Tese de doutorado em Engenharia de Produção, Universidade Federal de Santa Catarina.
- 11 BEN-AKIVA, Moshe.; LERMAN, S. R. Discrete choice analysis, **The MIT Press**, 1989.
- 12 _____.; BOCCARA, Bruno. Discrete choice models with latent choice sets. **Research in Marketing**, v. 12, p. 9-24, 1995.
- 13 BETTMAN, J. An information processing theory of consumer choice. **Reading, MA : Addison-Wesley**, 1979.

- 14 BOSE, R. C.; NAIR, K. R. Partially balanced incomplete block designs. **Sankhyā**, 4, 337-372, 1939.
- 15 BOX, G. E.; HUNTER, W.G.; HUNTER, J. Stuart. **Statistics for experimenters**. New York : John Wiley, 1978. 653 p.
- 16 CHAPMAN, R. G.; STAELIN, R. Exploiting rank ordered choice set data within the stochastic utility model. **Journal of Marketing Research**. n. 19, p. 288-301, 1982.
- 17 COBRA, Marcos; RANGEL, Alexandre. **Serviços ao cliente: uma estratégia competitiva**. 2 ed. São Paulo : MB, 1993.
- 18 COCHRAN, William G.; COX, Gerturude M. **Diseños Experimentales**. México : Trillas, 1978.
- 19 CONSTANTINO, Ademir Aparecido. **Otimização de escala de trabalho para condutores de trem : seqüenciamento de tarefas e alocação baseada em preferência declarada**. Florianópolis, 1997. 142p. Tese (Doutorado em Engenharia de Produção) - Curso de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina.
- 20 DALY, Andrew. Estimating "tree" logit models. **Transportation Research Record**, Great Britain, v. 21B, n. 4, p. 251-267, 1987.
- 21 DALY, Andrew. Applicability of desegregate behavioral modeling. **Transportation Research Record**, Great Britain, v. 16A, p. 363-370, 1984.
- 22 _____; ZACHARY, S. Improved multiple choice models. in : **Determinants of Travel Choice** (Edited by D. ^a Hensher and M. Q. Dalvi). Saxon House, Sussex, 1978.
- 23 DOMENCICH, T.; McFADDEN, L. Urban travel demand. **A Behavioral Analysis**, North Holland, Amsterdam, 1975.
- 24 EFRÓN, A. J. **Indústria hoteleira em Balneário Camburiú : uma visão através de modelos de preferência declarada**. Florianópolis, 1995. Dissertação de mestrado em Engenharia de Produção – Universidade Federal de Santa Catarina.
- 25 **Electric Power Research Institute**. Methodology for predicting the demand for new electricity —Using goods. EA-593, Project 488, Final report. Electric Power Research Institute, Palo Alto, California, 1977.
- 26 FINNEY, D. J. The fractional replication of factorial arrangements. **Ann. Eugen.** n. 12, p. 291-301, 1945.

- 27 FISHER, R. A. **An examination of the different possible solutions of a problem in incomplete blocks.** *Annals of Eugenics*, London, p. 52-75, 1940.
- 28 _____. **The design of experiments.** Oliver and Boyd, Edinburg, 4a. ed, 1947.
- 29 _____. YATES, F. **Statistical tables.** 3rd ed. Edinburgh : Oliver and Boyd, 1948.
- 30 FREITAS, Ana Augusta Ferreira. **Modelagem comportamental dos decisores através de técnicas de preferência declarada: uma aplicação no setor imobiliário de Florianópolis, SC.** Florianópolis, 1995. Dissertação de mestrado em Engenharia de Produção – Universidade Federal de Santa Catarina.
- 31 GAUDRY, Marc J. I.; WILLS, Michael J. Estimating the functional form of travel demand models. **Centre de recherche sur les Transports, Université de Montréal**, Montreal, p. 257-289, Received 22 Nov. 1978.
- 32 _____. JARA-DIAZ, Sérgio R.; ORTÚZAR, Juan de Dios. Value of time sensitivity to model specification. **Transportation Research : an international journal**, New York, v. 23B, n. 2, p. 151-158, April 1989.
- 33 GENSCH, Dennis H.; SOOFI, Ehasan S. Information-theoretic estimation of individual consideration set. **Research in Marketing**, v. 12, p. 25-38, 1995.
- 34 GONÇALVES, Natália M. **Economias de escala em uma linha de ônibus urbano : o enfoque micro - analítico.** Florianópolis, 1995. Dissertação de mestrado em Engenharia de Produção – Universidade Federal de Santa Catarina.
- 35 GREEN, P. E. On the design of choice experiments involving multifactor alternatives. **Journal of Consumer Research**, v. 1, n. 2, p. 61-68, 1974.
- 36 GREEN, P. E.; RAO, V. R. Conjoint measurement for quantifying judgmental data. **Journal of Marketing Research**, v. 8, n. 3, p. 355-363, 1971.
- 37 _____. WIND, Y. New way to measure consumers judgments. **Harvard Business Review**, v. 53, n. 4, p. 107-117, 1975.
- 38 _____. SRINIVASAN, V. Conjoint analysis in consumer research: issues and outlook. **Journal of Consumer Research**, v. 5, n. 2, p. 102-123, 1978.
- 39 _____. TULL, D. S. **Research for marketing decisions.** 4 ed. Englewood Cliffs, Prentice Hall, 1978.
- 40 HAHN, G. J.; SHAPIRO, S. S. A catalogue and computer programmer for design and analysis of orthogonal symmetric and asymmetric fractional experiments. **General Electric Research and development Center Report**. N. 66-C-165, Schenectady, New York, 1966.
- 41 HARSHBARGER, R. Triple rectangular lattices. **Biometrics**. 5, 1-13, 1949.

- 42 HAUSMAN, J.; McFADDEN, D. Specification tests for the multinomial logit models. **Econometrics** n. 52, p. 1219-1240, 1984.
- 43 HENSHER, David A. et. all. Urban toll-ways and the valuation of travel time saving. **The Economic Record** 66(193), p. 146-156, 1989.
- 44 _____.; BARNARD, P. O. The orthogonality issue in stated choice design. In: Fischer M. Nijkamp, P. & Papageorgiou, **Spatial Choices and Processes**. p. 265-278, North Holland, Amsterdam, 1990.
- 45 _____. Stated preference analysis of travel choice : the state of practice. **Transportation**, Netherlands, n. 21, p. 107-133, 1994.
- 46 HERNSTEIN, I. N., **Tópicos de álgebra**. tradução de Adalberto P. Bergamasco e L. H. Jacy Monteiro, São Paulo: Polígono, 1970.
- 47 HOFFMAN, P. J. The paramorphic representation of clinical judgment. **Psychological Bulletin**. n. 57, p. 116-131, 1960.
- 48 _____.; SLOVIC, P.; RORER, L. G. An analysis-of-variance model for the assessment of configure cue utilization in clinical judgment. **Psychological Bulletin**. n. 69, p. 338-349, 1968.
- 49 HOROWITZ, Joel L.; LOUVIERE, Jordon J. What is the role of consideration sets in choice modeling? **Research in Marketing**, v. 12, p. 39-54, 1995.
- 50 JOHNSON, R. M. Trade-off analysis of consumer values. **Journal of Marketing Research**, v. 11, n. 2, p. 121-127, 1974.
- 51 JONES, P. **An overview of stated preference techniques**, PTRC : Course: Introduction to Stated Preference Techniques, 1991. p. 1-14.
- 52 KEENEY, R. L. **Value-focused thinking: a path to creative decision making**. Cambridge : Harvard Univ. Press, 1992.
- 53 KEMPTHORNE, Oscar. **The design and analysis of experiments**. New York : John Willy, 1967.
- 54 KITAMURA, R.; LAM, T. A model of constrained binary choice. In: J. Volmuller and R. Hamerslag (eds.), **Proceedings of the Ninth International Symposium on Transportation and Traffic Theory**, p. 493-512. Utrecht, The Netherlands : VNU Science Press. 1984.
- 55 KNAPP, Rebecca G.; MILLER III, M. Clinton. **Clinical epidemiology and biostatistics**. Williams & Wilkns, Baltimore, 1992, 435p.

- 56 KROES, Eric P.; SHELDON, Robert J. Stated preference methods. An introduction. **Journal of Transport Economics and Policy**, London, v. XXII, n. 1, p. 11-25, January 1988.
- 57 LANCASTER, K. A new approach to consumer theory. **Journal of Political Economy**, n. 74, p. 132-157, 1966.
- 58 LAPERSONNE, Eric; LAURENT, Gilles; LE GOFF, Jean-Jacques. Consideration sets of size one: an empirical investigation of automobile purchases. **Research in Marketing**, v. 12, p. 55-66, 1995.
- 59 LEVITT, Theodore. **A imaginação de marketing**. São Paulo : Atlas, 1990.
- 60 LOUVIERE, Jordan J. Conjoint analysis modeling of stated preferences : a review of theory, methods, recent developments and external validity. **Journal of Transport Economics and Policy**, London, v. XXII, n. 1, p. 93-119, January 1988.
- 61 ____; MEYER, Robert. A composite attitude-behavior model of traveler decision Making. **Transportation Research**, 15B, p. 411-420, December, 1981.
- 62 ____; HENSHER, D. A. Using discrete choice models with experiments design data to forecast consumer demand for a unique cultural event. **Journal of Consumer Research**. v. 3, p. 348-381, 1983.
- 63 ____; GAETH, Gary J. Decomposing the determinants of retail facility choice using the method of hierarchical information integration: a supermarket illustration. **Journal of Retailing**, Aiwa City, Iowa, v. 63, n. 1, p. 25-47, Spring 1987.
- 64 _____. Conjoint analysis. In: BAGOZZI, R. P. **Methods of marketing research**. Cambridge, USA : Blackwell Publishes, p. 223-259, 1994.
- 65 LUCE. R.; SUPPES, P. Preference, Utility ad subjective probability. In **Handbook of Mathematical Psychology**. v. 3, R. Luce, and E. Galanter, eds. Wiley, New York, 1965.
- 66 MALHOTRA, N. K. Structural reliability and stability of non-metric conjoint analysis. **Journal of Marketing Research**, v. 19, n. 2, p. 199-207, 1982.
- 67 MALHOTRA, N. K. The use of linear logit models in marketing research. **Journal of Marketing Research**, v. 21, p. 20-31, February, 1984.
- 68 MANDEL, J. Chain block designs with two-way elimination of heterogeneity. **Biometrics**. n. 10, p. 251-272, 1954.
- 69 MANN, H. B. **Analysis and design of experiments**. New York : Dover Publications, 1949.

- 70 MASON, C. H.; PERREAULT, W. D. Collinearity, power, and interpretation of multiple regression analysis. **Journal of Marketing Research XXVIII**, p. 268-280, 1991.
- 71 MAY, Gabriela de Oliveira. **Uma aplicação das técnicas de preferência declarada ao setor hoteleiro de Florianópolis**. Florianópolis, 1996. Dissertação de mestrado em Engenharia de Produção – Universidade Federal de Santa Catarina.
- 72 McFADDEN, Daniel L. Conditional logit analysis of qualitative choice behavior. **Frontiers in Econometrics**, New York : Academic Press. p. 105-142. 1974.
- 73 _____. Modeling the choice of residential location. **Transportation Research Record** n.673, p. 72-77, 1978.
- 74 _____. Quantitative models for analyzing travel behavior of individuals, in **Behavioral Travel Modeling** (Edited by D. Hensher and P.R. Stopher). Croom Helm, London, 1979.
- 75 _____. Econometric analysis of qualitative response models. In: **Handbook Of Econometrics**, v. 2. Massachusetts Institute of Technology : Z. Griliches, p. 1395-1457, 1984.
- 76 _____. TYE, William B.; TRAIN, Kenneth. An application of diagnostic testes for the independence from irrelevant alternatives property of the multinomial logit model. **Transportation Research Record** n. 637, p. 39-46, 1977.
- 77 _____. Econometric models of probabilistic choice. In : **Structural Analysis Of Discrete Data with Econometric Applications** (Edited by D. McFadden and C. F. Manski). MIT Press, Cambridge, Massachusetts, 1981.
- 78 McLEAN, R.; ANDERSON, V. L. **Applied factorial and fractional designs**, New York : Marcol Derbei, Inc, 1994.
- 79 MEAD, R. **The design of experiments**. 620p. Cambridge : Cambridge Univ., 2nd ed., 1992.
- 80 MILLER, G. A. **The magic number seven plus or minus two: some limits on the capacity for processing information**. Psych. Rev. 63, p. 81-97, 1956.
- 81 MONTGOMERY, Douglas C. **Design and analysis of experiments**. 2nd ed. New York : J. Wiley, 1984.
- 82 MORIKAWA, T. **Incorporating stated preference data in travel demand analysis**, Ph.D. Thesis, MIT, 1989.

- 83 NANSKI, C. The structure of random utility models. **Theory e Decision**, v. 8, p. 229-254, 1977.
- 84 NOVAES, A. G. N.; CARVALHO, M. C. Market share analysis of transport services with stated preference data. **Coletânea Politécnica – Revista Brasileira de Tecnologia e Ciência**, USP : São Paulo, v. 1. n. 1. p. 78-91, Maio/agosto 1996.
- 85 OLIVEIRA, Adelize Generini de,. **Técnicas Avançadas de programação em Delphi**. Florianópolis : Bookstore, 1996.
- 86 OPPEWAL, Harmen; LOUVIERE, Jordan J.; TIMMERMANS, Harry J. P. Modeling hierarchical conjoint processes with integrated choice experiments. **Journal of Marketing Research**. v. XXXI, p. 92-104, 1995.
- 87 ORTUZAR, Juan de Dios; WILLUMSEN, Luis G. **Modeling transport**. Chichester England : John Wiley & Sons, 1990.
- 88 _____. _____. 2 ed. 1994.
- 89 PEKELMAN, Dov; SEM, Subrata K. Measurement and estimation of conjoint utility functions. **Journal of Consumer Research**, [S.l.], v. 5, p. 263-271, March 1979.
- 90 PERSON, E. S.; HARTLEY, H. O. **Biometrika tables for statisticians**. v. 1, Cambridge : University Press, 1970. 270 p.
- 91 PIMENTEL, GOMES, F., **Curso de Estatística Experimental**, 12 ed. rev. amp., Piracicaba: Nobel, 1987.
- 92 ROSS, Phillip J. **Aplicações das técnicas de Taguchi na engenharia da qualidade**. São Paulo : Makron, 1991.
- 93 SHELDON, R. **Stated Preference: Design Issues**, PTRC: Course: Introduction to Stated Preference Techniques, 1991.
- 94 SHOCKER, A. D.; BEN-AKIVA, M.; BOCCARA, B.; NEDUNGADI, P. Consideration set influences on consumer decision-making and choice : issues, models, and suggestions. **Marketing Letters** v. 2, p. 181-198, 1991.
- 95 SWAIT, J. **Probabilistic choice set formation in transportation demand models**. Unpublished Ph.D. Dissertation, MIT, Cambridge, MA, 1984.
- 96 TAGUCHI, Genichi. **System of experimental design : engineering methods to optimize quality and minimize costs**. New York : Kaus I. P., 1988.
- 97 VIEIRA, Hélio Flávio. **Uma visão empresarial do processo de exportação de produtos containerizados catarinense e análise do nível de serviço**.

Florianópolis, 1996. Dissertação de mestrado em Engenharia de Produção – Universidade Federal de Santa Catarina.

- 98 VINCKE, Philippe. **Multicriteria decision-aid**. New York : Willey, 1992.
- 99 WARDMAN, Mark. A comparison of revealed preference and stated preference models of travel behavior. **Journal of Transport Economics and Policy**, London, v. XXII, n. 1, p. 71-91, January 1988.
- 100 WESTIN, R. D.; MANSKI, C. F. Theoretical and conceptual developments in demand modeling, in **Behavioral Travel Modeling** (Edited by D. A. Hensher and P. R. Stopher), Croom Helm, London, 1979.
- 101 WILLIAMS, H. C. W. L. On the formation of travel demand models and economic measures of user benefit. **Environment and Planning**. 9A., p. 285-344, 1977.
- 102 YATES, Frank. Incomplete randomized blocks. In: **Experimental design** : selected papers of Frank Yates. London : Charles Griffin, 1936.
- 103 _____. A new method of arranging variety trials involving a large number of varieties, **Jour. Agr. Sci.** 26, 424-45, 1936.

ANEXOS

Anexo	Assunto	Página
1	TABELA DE EXPERIMENTOS FATORIAIS	142
2	ARRANJOS ORTOGONAIS	143
3	ARRANJOS FATORIAIS	149
4	CONJUNTO DE ESCOLHA COM PROBABILIDADE CONDICIONAL	155
5	EXEMPLOS DE UTILIZAÇÃO DAS TABELAS	163
6	ASPECTOS COMPUTACIONAIS	166
7	QUESTIONÁRIO PARA ENTREVISTAS	169
8	EXEMPLOS DE CARTÕES	173
9	RELATÓRIO DAS ENTREVISTAS	177

Anexo 1

TABELA DE EXPERIMENTOS FATORIAIS

O Esquema fatorial a ser adotado depende de alguns fatores:

- do pressuposto da ortogonalidade entre os atributos, o que possibilitará a montagem de delineamento de arranjos ortogonais;
- do número de atributos e do número de níveis em cada atributo.

TABELA DE EXPERIMENTOS FATORIAIS

Nr	Atributos	NÍVEIS			Ensaio Fatorial	Anexo 2 Tabela	Anexo 3 Ensaio
		2	3	4			
01	3	3	-	-	2^3	A1	3.1
02	3	2	*	1	$2^2 \times 4$	A2	-
03	3	1	*	2	2×4^2	A7	-
04	3	-	3	-	3^3	A5	4.1
05	3	-	-	3	4^3	A6	4.3
06	4	4	-	-	2^4	A1	1.1 – 2.1 – 3.2
07	4	3	*	1	$2^3 \times 4$	A2	-
08	4	2	*	2	$2^2 \times 4^2$	A7	-
09	4	1	3	-	2×3^3	A10	-
10	4	1	-	3	2×4^3	A8	-
11	4	-	4	-	3^4	A5	2.14 – 2.15 – 4.1
12	4	-	-	4	4^4	A6	4.3
13	5	5	-	-	2^5	A1	2.2 – 2.3
14	5	4	*	1	$2^4 \times 4$	A2	-
15	5	3	*	2	$2^3 \times 4^2$	A7	-
16	5	2	*	3	$2^2 \times 4^3$	A8	-
17	5	1	4	-	2×3^4	A10	-
18	5	1	-	4	2×4^4	A9	-
19	5	-	5	-	3^5	A5	4.1
20	5	-	-	5	4^5	A6	4.3
21	6	6	-	-	2^6	A1 – A3	2.5 ... 2.8 – 3.3
22	6	5	*	1	$2^5 \times 4$	A4	-
23	6	4	*	2	$2^4 \times 4^2$	A7	-
24	6	3	*	3	$2^3 \times 4^3$	A8	-
25	6	2	4	-	$2^2 \times 3^4$		-
26	6	2	-	4	$2^2 \times 4^4$	A9	-
27	6	1	5	-	2×3^5	A10	-
28	7	7	-	-	2^7	A1 – A3	2.9 ... 2.13
29	7	6	*	1	$2^6 \times 4$	A4	-

* Pode ser designado atributos com 3 níveis reduzindo os de 4 níveis;

A divisão do conjunto de alternativas em blocos, ou seja, conjunto de escolha com probabilidade condicional, baseadas nas técnicas de blocos incompletos, deverão se feitas de acordo com a tabela A4.1 do anexo 4.

Anexo 2

ARRANJOS ORTOGONAIS

Para a utilização das tabelas, em experimentos de Preferência Declarada, devem ser observadas as seguintes regras e procedimentos:

- 1) Cada letra (coluna) da tabela corresponde a um atributo com 2, 3 ou 4 níveis, conforme o caso e cada linha corresponde às alternativas que são formadas pelos elementos (níveis) das colunas;
- 2) Existe uma correspondência entre os parâmetros da função utilidade (FU) com as colunas (letras) referentes aos atributos;
- 3) Para atributos com 3 ou 4 níveis:
 - Caso seja discreto: utilizar dois parâmetros correspondentes às duas colunas do atributo, cada coluna com dois níveis; utilizar o quarto nível, no caso de atributo com 3 níveis, com a repetição de um dos 3 níveis, o que achar mais conveniente;
 - Caso seja contínuo: utilizar um parâmetro correspondente à coluna do atributo com 3 ou 4 níveis;
 - Caso existem combinações de atributos com níveis discretos e contínuos, deve proceder conforme o caso;
- 4) Caso a quantidade de atributos seja menor que a quantidade de colunas, em um dado arranjo, basta excluir os parâmetros da FU e os elementos da coluna correspondentes, no arranjo. O número de alternativas não é alterado.
- 5) Para atributos com dois níveis, os procedimentos serão iguais tanto para níveis discretos quanto contínuos;
- 6) Todas as tabelas de arranjos fatoriais foram montadas baseadas nos Arranjos Ortogonais de Taguchi (ROSS, 1991).

1) Ensaios fatoriais: 2^3 ; 2^4 ; 2^5 ; 2^6 e 2^7

$$FU = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$$

Tabela A1 - Ensaios fatoriais com atributos de 2 níveis: L8

Alternativas	A	B	C	D	E	F	G
	β_1	β_2	β_3	β_4	β_5	β_6	β_7
01	0	0	0	0	0	0	0
02	0	0	1	1	0	1	1
03	0	1	0	1	1	0	1
04	0	1	1	0	1	1	0
05	1	0	0	1	1	1	0
06	1	0	1	0	1	0	1
07	1	1	0	0	0	1	1
08	1	1	1	1	0	0	0

2) Ensaios fatoriais: 4×2^2 ; 4×2^3 e 4×2^4

Caso o atributo com 4 níveis seja:

Discreto:
$$FU = \beta_{11} X_{11} + \beta_{12} X_{12} + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

Contínuo:
$$FU = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

Tabela A2 - Ensaios fatoriais com atributos de 4 e 2 níveis: L8

Alternativas	A			B	C	D	E
	β_{11}	β_{12}	β_1	β_2	β_3	β_4	β_5
01	0	0	0	0	0	0	0
02	0	0	0	1	1	1	1
03	0	1	1	0	1	0	1
04	0	1	1	1	0	1	0
05	1	0	2	0	1	1	0
06	1	0	2	1	0	0	1
07	1	1	3	0	0	1	1
08	1	1	3	1	1	0	0

3) Ensaios fatoriais: 2^6 e 2^7

$$FU = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$$

É conveniente utilizar este arranjo quando se tem 6 ou 7 atributos.

Tabela A3 - Ensaios fatoriais com 2 níveis: L12

Alternativas	A	B	C	D	E	F	G
	β_1	β_2	β_3	β_4	β_5	β_6	β_7
01	0	0	0	0	0	0	0
02	1	1	1	1	1	1	1
03	1	0	0	0	1	1	1
04	1	0	1	1	0	0	1
05	1	1	0	1	0	1	0
06	0	1	1	0	1	0	0
07	0	0	1	1	0	1	0
08	1	1	1	0	0	0	1
09	1	1	0	1	1	0	0
10	0	0	0	1	1	0	1
11	0	1	0	0	0	1	1
12	0	0	1	0	1	1	0
	5	6	7	8	9	10	11

4) Ensaios fatoriais: 4×2^5 e 4×2^6

Caso o atributo com 4 níveis seja:

Discreto: $FU = \beta_{11} X_{11} + \beta_{12} X_{12} + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$

Contínuo: $FU = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$

Tabela A4 – Ensaios fatoriais com 4 e 2 níveis: L12

Alternativas	A		C	D	E	F	G	H
	β_{11}	β_{12}	β_1	β_2	β_3	β_4	β_5	β_7
01	0	0	0	0	0	0	0	0
02	0	0	0	0	0	0	1	1
03	0	0	0	1	1	1	0	0
04	0	1	1	0	1	1	0	1
05	0	1	1	1	0	1	1	0
06	0	1	1	1	1	0	1	0
07	1	0	2	1	1	0	0	1
08	1	0	2	1	0	1	1	0
09	1	0	2	0	1	1	0	1
10	1	1	3	1	0	0	0	1
11	1	1	3	0	1	0	1	0
12	1	1	3	0	0	1	0	0

5) Ensaios fatoriais: 3^3 ; 3^4 e 2^5

Caso os atributos com 3 níveis sejam: Discretos:

$$FU = \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{21}X_{21} + \beta_{22}X_{22} + \beta_{31}X_{31} + \beta_{32}X_{32} + \beta_{41}X_{41} + \beta_{42}X_{42}$$

Contínuos: $FU = \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$

Tabela A5 – Ensaios fatoriais com atributos de 3 níveis: L9

Alternativas	A			B			C			D		
	β_{11}	β_{12}	β_1	β_{21}	β_{22}	β_2	β_{31}	β_{32}	β_3	β_{41}	β_{42}	β_4
01	0	0	1	0	0	1	0	0	1	0	0	1
02	0	0	1	0	1	2	0	1	2	0	1	2
03	0	0	1	1	0	3	1	0	3	1	0	3
04	0	1	2	0	0	1	0	1	2	1	0	3
05	0	1	2	0	1	2	1	0	3	0	0	1
06	0	1	2	1	0	3	0	0	1	0	1	2
07	1	0	3	0	0	1	1	0	3	0	1	2
08	1	0	3	0	1	2	0	0	1	1	0	3
09	1	0	3	1	0	3	0	1	2	0	0	1

6) Ensaios fatoriais: 4^3 ; 4^4 e 4^5

Caso os atributos com 4 níveis sejam: Discretos:

$$FU = \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{21}X_{21} + \beta_{22}X_{22} + \beta_{31}X_{31} + \beta_{32}X_{32} + \beta_{41}X_{41} + \beta_{42}X_{42} + \beta_{51}X_{51} + \beta_{52}X_{52}$$

Contínuos: $FU = \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5$

Tabela A6 – Ensaios fatoriais com atributos de 4 níveis: L16

Alternativas	Atributo A			Atributo B			Atributo C			Atributo D			Atributo E		
	β_{11}	β_{12}	β_1	β_{21}	β_{22}	β_2	β_{31}	β_{32}	β_3	β_{41}	β_{42}	β_4	β_{51}	β_{52}	β_5
01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1
03	0	0	0	1	0	2	1	0	2	1	0	2	1	0	2
04	0	0	0	1	1	3	1	1	3	1	1	3	1	1	3
05	0	1	1	0	0	0	0	1	1	1	0	2	1	1	3
06	0	1	1	0	1	1	0	0	0	1	1	3	1	0	2
07	0	1	1	1	0	2	1	1	3	0	0	0	0	1	1
08	0	1	1	1	1	3	1	0	2	0	1	1	0	0	0
09	1	0	2	0	0	0	1	0	2	1	1	3	0	1	1
10	1	0	2	0	1	1	1	1	3	1	0	2	0	0	0
11	1	0	2	1	0	2	0	0	0	0	1	1	1	1	3
12	1	0	2	1	1	3	0	1	1	0	0	0	1	0	2
13	1	1	3	0	0	0	1	1	3	0	1	1	1	0	2
14	1	1	3	0	1	1	1	0	2	0	0	0	1	1	3
15	1	1	3	1	0	2	0	1	1	1	1	3	0	0	0
16	1	1	3	1	1	3	0	0	0	1	0	2	0	1	1

7) Ensaios fatoriais: $4^2 \times 2$; $4^2 \times 2^2$; $4^2 \times 2^3$; $4^2 \times 2^4$

Caso os atributos com 4 níveis sejam:

Discretos: $FU = \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{21}X_{21} + \beta_{22}X_{22} + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$

Contínuos: $FU = \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$

Tabela A7 – Ensaios fatoriais com atributos de 4 e 2 níveis: L16

Alternativas	A			B			C	D	E	F
	β_{11}	β_{12}	β_1	β_{21}	β_{22}	β_2	β_3	β_4	β_5	β_6
01	0	0	0	0	0	0	0	0	0	0
02	0	0	0	0	1	1	0	0	1	1
03	0	0	0	1	0	2	1	1	0	0
04	0	0	0	1	1	3	1	1	1	1
05	0	1	1	0	0	0	0	1	0	1
06	0	1	1	0	1	1	0	1	1	0
07	0	1	1	1	0	2	1	0	0	1
08	0	1	1	1	1	3	1	0	1	0
09	1	0	2	0	0	0	1	1	1	0
10	1	0	2	0	1	1	1	1	0	1
11	1	0	2	1	0	2	0	0	1	0
12	1	0	2	1	1	3	0	0	0	1
13	1	1	3	0	0	0	1	0	1	1
14	1	1	3	0	1	1	1	0	0	0
15	1	1	3	1	0	2	0	1	1	1
16	1	1	3	1	1	3	0	1	0	0

8) Ensaios fatoriais: $4^3 \times 2$; $4^3 \times 2^2$ e $4^3 \times 2^3$

Caso os atributos com 4 níveis sejam: Discretos:

$FU = \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{21}X_{21} + \beta_{22}X_{22} + \beta_{31}X_{31} + \beta_{32}X_{32} + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$

Contínuos: $FU = \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$

Tabela A8 – Ensaios fatoriais com atributos de 4 e 2 níveis: L16

Alt.	A			B			C			D	E	F
	β_{11}	β_{12}	β_1	β_{21}	β_{22}	β_2	β_{31}	β_{32}	β_3	β_4	β_5	β_6
01	0	0	0	0	0	0	0	0	0	0	0	0
02	0	0	0	0	1	1	0	1	1	0	0	1
03	0	0	0	1	0	2	1	0	2	1	1	0
04	0	0	0	1	1	3	1	1	3	1	1	1
05	0	1	1	0	0	0	0	1	1	1	1	0
06	0	1	1	0	1	1	0	0	0	1	1	1
07	0	1	1	1	0	2	1	1	3	0	0	0
08	0	1	1	1	1	3	1	0	2	0	0	1
09	1	0	2	0	0	0	1	0	2	0	1	1
10	1	0	2	0	1	1	1	1	3	0	1	0
11	1	0	2	1	0	2	0	0	0	1	0	1
12	1	0	2	1	1	3	0	1	1	1	0	0
13	1	1	3	0	0	0	1	1	3	1	0	1
14	1	1	3	0	1	1	1	0	2	1	0	0
15	1	1	3	1	0	2	0	1	1	0	1	1
16	1	1	3	1	1	3	0	0	0	0	1	0

9) Ensaios fatoriais: $4^4 \times 2$ e $4^4 \times 2^2$

Caso os atributos com 4 níveis sejam: Discretos:

$$FU = \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{21}X_{21} + \beta_{22}X_{22} + \beta_{31}X_{31} + \beta_{32}X_{32} + \beta_{41}X_{41} + \beta_{42}X_{42} + \beta_5X_5 + \beta_6X_6$$

$$\text{Contínuos: } FU = \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$$

Tabela A9 – Ensaios fatoriais com atributos de 4 e 2 níveis: L16

Alt.	A			B			C			D			E	F
	β_{11}	β_{12}	β_1	β_{21}	β_{22}	β_2	β_{31}	β_{32}	β_3	β_{41}	β_{42}	β_4	β_5	β_6
01	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02	0	0	0	0	1	1	0	1	1	0	1	1	0	1
03	0	0	0	1	0	2	1	0	2	1	0	2	1	0
04	0	0	0	1	1	3	1	1	3	1	1	3	1	1
05	0	1	1	0	0	0	0	1	1	1	0	2	1	1
06	0	1	1	0	1	1	0	0	0	1	1	3	1	0
07	0	1	1	1	0	2	1	1	3	0	0	0	0	1
08	0	1	1	1	1	3	1	0	2	0	1	1	0	0
09	1	0	2	0	0	0	1	0	2	1	1	3	0	1
10	1	0	2	0	1	1	1	1	3	1	0	2	0	0
11	1	0	2	1	0	2	0	0	0	0	1	1	1	1
12	1	0	2	1	1	3	0	1	1	0	0	0	1	0
13	1	1	3	0	0	0	1	1	3	0	1	1	1	0
14	1	1	3	0	1	1	1	0	2	0	0	0	1	1
15	1	1	3	1	0	2	0	1	1	1	1	3	0	0
16	1	1	3	1	1	3	0	0	0	1	0	2	0	1

10) Ensaios fatoriais: $3^4 \times 2$ e $3^5 \times 2$

Caso os atributos com 3 níveis sejam: Discretos:

$$FU = \beta_{11}X_{11} + \beta_{12}X_{12} + \beta_{21}X_{21} + \beta_{22}X_{22} + \beta_{31}X_{31} + \beta_{32}X_{32} + \beta_{41}X_{41} + \beta_{42}X_{42} + \beta_{51}X_{51} + \beta_{52}X_{52} + \beta_6X_6$$

$$\text{Contínuos: } FU = \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6$$

Tabela A10 – Ensaios fatoriais com atributos de 3 e 2 níveis: L18

Alt.	A			B			C			D			E			F
	β_{11}	β_{12}	β_1	β_{21}	β_{22}	β_2	β_{31}	β_{32}	β_3	β_{41}	β_{42}	β_4	β_{51}	β_{52}	β_5	β_6
01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1	0
03	0	0	0	1	0	2	1	0	2	1	0	2	1	0	2	0
04	0	1	1	0	0	0	0	0	0	0	1	1	0	1	1	0
05	0	1	1	0	1	1	0	1	1	1	0	2	1	0	2	0
06	0	1	1	1	0	2	1	0	2	0	0	0	0	0	0	0
07	1	0	2	0	0	0	0	1	1	0	0	0	1	0	2	0
08	1	0	2	0	1	1	1	0	2	0	1	1	0	0	0	0
09	1	0	2	1	0	2	0	0	0	1	0	2	0	1	1	0
10	0	0	0	0	0	0	1	0	2	1	0	2	0	1	1	1
11	0	0	0	0	1	1	0	0	0	0	0	0	1	0	2	1
12	0	0	0	1	0	2	0	1	1	0	1	1	0	0	0	1
13	0	1	1	0	0	0	0	1	1	1	0	2	0	0	0	1
14	0	1	1	0	1	1	1	0	2	0	0	0	0	1	1	1
15	0	1	1	1	0	2	0	0	0	0	1	1	1	0	2	1
16	1	0	2	0	0	0	1	0	2	0	1	1	1	0	2	1
17	1	0	2	0	1	1	0	0	0	1	0	2	0	0	0	1
18	1	0	2	1	0	2	0	1	1	0	0	0	0	1	1	1

Anexo 3

ARRANJOS FATORIAIS

Os arranjos fatoriais constante deste anexo foram montados com base nos arranjos desenvolvidos por COCHRAN e COX (1978), adaptados para sua utilização nas Técnicas de Preferência Declarada.

Limitações imposta pela Técnica de Preferência Declarada (de acordo com a Tese):

- Atributos:
 - Até 7 com 2 níveis;
 - Até 5 com 3 ou níveis;
 - Até 6 mistos com 2 e 3 ou 4 níveis;
- Total de alternativas:
 - Até 18 alternativas;
- Blocos:
 - Até 20 blocos;
- Alternativa em cada bloco:
 - De 2 a 9 alternativas;

1. Experimentos Fatoriais com Repetição Simples

Ensaio 1.1 – Fatorial 2^4 com blocos balanceados com 4 alternativas:

Repetição I

Blocos	Alternativas			
(1)	0000	1110	1101	0011
(2)	1100	0010	0001	1111
(3)	1000	0110	0101	1011
(4)	0100	1010	1001	0111

Repetição II:

Blocos	Alternativas			
(5)	0000	1011	0111	1100
(6)	0011	1000	0100	1111
(7)	0010	1001	0101	1110
(8)	0001	1010	0110	1101

2. Experimentos fatoriais com repetição fracionada (Fatorial fracionário e confundidos)

Freqüentemente uma só repetição de um experimento fatorial vai além dos recursos dos investigadores, ou dá mais precisão do que o necessário para a estimação dos efeitos principais (COCHRAN e COX, 1978).

O uso de experimentos em repetição fracionada foi proposto em 1945 por Finney¹. Desde então, os projetos foram utilizados em muitas aplicações, particularmente em desenvolvimento industriais. Sua principal atratividade é que permite incluir 5 ou mais fatores em um experimento de tamanho prático, de tal maneira que o investigador pode determinar rapidamente os efeitos dos fatores no resultado.

A diminuição do tamanho dos experimentos não se obtém sem sacrificar outros aspectos. Assim os aspectos sacrificados são algumas possíveis interações entre dois efeitos principais.

Ensaio 2.1 Fatorial 2^4 - 1/2 repetição ($v = 8$; $b = 1$; $k = 8$):

Bloco Alternativas:

(1) 0000 1100 1010 1001 0110 0101 0011 1111

Ensaio 2.2 Fatorial 2^5 - 1/4 repetição ($v = 8$; $b = 1$; $k = 8$):

Bloco Alternativas:

(1) 00000 11000 00110 10101 01101 10011 01011 11111

Ensaio 2.3 Fatorial 2^5 - 1/2 repetição ($v = 16$; $b = 4$; $k = 4$):

Blocos Alternativas:

(1) 00000 11000 10111 01111

(2) 10100 01100 00011 11011

(3) 10001 01001 00110 11111

(4) 10010 01001 00101 11101

¹ FINNEY, D. J. The fractional replication of factorial arrangements. *Ann. Eugen.* 12, 291-301, 1945.

Ensaio 2.4 Fatorial 2^5 - 1/2 repetição ($v = 16$; $b = 2$; $k = 8$):

Blocos Alternativas

- (1) 00000 11000 10111 01111 10100 01100 00011 11011
 (2) 10001 01001 00110 11111 10010 01001 00101 11101

Ensaio 2.5 Fatorial 2^6 - 1/8 repetição ($v = 8$; $b = 1$; $k = 8$):

Bloco Alternativas

- (1) 000000 101001 100110 011010 010101 111100 110011 001111

Ensaio 2.6 Fatorial 2^6 - 1/4 repetição ($v = 16$; $b = 4$; $k = 4$):

Bloco Alternativas

- (1) 000000 111010 110101 001111
 (2) 101100 100011 011001 010110
 (3) 110000 001010 000101 111111
 (4) 101001 100110 011100 010011

Ensaio 2.7 Fatorial 2^6 - 1/4 repetição ($v = 16$; $b = 2$; $k = 8$):

Bloco Alternativas

- (1) 000000 111010 110101 001111 101100 100011 011001 010110
 (2) 110000 001010 000101 111111 101001 100110 011100 010011

Ensaio 2.8 Fatorial 2^6 - 1/2 repetição ($v = 32$; $b = 8$; $k = 4$):

Bloco Alternativas

- (1) 000000 110011 101110 011101
 (2) 110000 000011 101101 011110
 (3) 101000 000110 110101 011011
 (4) 011000 000101 101011 110110
 (5) 100010 010001 001100 111111
 (6) 100001 010010 111100 001111
 (7) 100100 001010 111001 010111
 (8) 010100 001001 111010 100111

Ensaio 2.9 Fatorial 2^7 - 1/16 repetição ($v = 8$; $b = 1$; $k = 8$):

Bloco Alternativas

- (1) 0000000 1111000 1100110 1010011 1001101 0110101 0101011 0011110

Ensaio 2.10 Fatorial 2^7 - 1/2 repetição ($v = 16$; $b = 4$; $k = 4$):

Bloco Alternativas

- | | | | | |
|-----|---------|---------|---------|---------|
| (1) | 0000000 | 0000111 | 1111000 | 1111111 |
| (2) | 1100001 | 0011001 | 1100110 | 0011110 |
| (3) | 1010010 | 0101010 | 1010101 | 0101101 |
| (4) | 1001100 | 0110100 | 1001011 | 0110011 |

Ensaio 2.11 Fatorial 2^7 - 1/8 repetição ($v = 16$; $b = 2$; $k = 8$):

Bloco Alternativas

- | | | | | | | | | |
|-----|---------|---------|---------|---------|---------|---------|---------|---------|
| (1) | 0000000 | 1111000 | 1100110 | 1010101 | 1001011 | 0110011 | 0101101 | 0011110 |
| (2) | 1100001 | 1010010 | 1001100 | 0110100 | 0101010 | 0011001 | 0000111 | 1111111 |

Ensaio 2.12 Fatorial 2^7 - 1/4 repetição ($v = 32$; $b = 8$; $k = 4$):

Bloco Alternativas

- | | | | | |
|-----|---------|---------|---------|---------|
| (1) | 0000000 | 0001111 | 1111010 | 1110101 |
| (2) | 0001100 | 0000011 | 1111001 | 1110110 |
| (3) | 1100000 | 0011010 | 0010101 | 1101111 |
| (4) | 0011001 | 0010110 | 1101100 | 1100011 |
| (5) | 1010000 | 0101010 | 0100101 | 1011111 |
| (6) | 0101001 | 0100110 | 1011100 | 1010011 |
| (7) | 0110000 | 1001010 | 1000101 | 0111111 |
| (8) | 1001001 | 1000110 | 0110011 | 0111100 |

Ensaio 2.13 Fatorial 2^7 - 1/2 repetição ($v = 64$; $b = 16$; $k = 4$):

Bloco Alternativas

- | | | | | |
|------|---------|---------|---------|---------|
| (1) | 0000000 | 1111000 | 0001111 | 1110111 |
| (2) | 1100000 | 0011000 | 1101111 | 0010111 |
| (3) | 1010000 | 0101000 | 1011111 | 0100111 |
| (4) | 0110000 | 1001000 | 0111111 | 1000111 |
| (5) | 1000100 | 0111100 | 1001011 | 0110011 |
| (6) | 0100100 | 1011100 | 0101011 | 1010011 |
| (7) | 0010100 | 1101100 | 0011011 | 1100011 |
| (8) | 1110100 | 0001100 | 1111011 | 0000011 |
| (9) | 1000010 | 0111010 | 1001101 | 0110101 |
| (10) | 0100010 | 1011010 | 0101101 | 1010101 |
| (11) | 0010010 | 1101010 | 0011101 | 1100101 |
| (12) | 1110010 | 0001010 | 1111101 | 0000101 |
| (13) | 0000110 | 1111110 | 0001001 | 1110001 |
| (14) | 1100110 | 0011110 | 1101001 | 0010001 |
| (15) | 1010110 | 0101110 | 1011001 | 0100001 |
| (16) | 0110110 | 1001110 | 0111001 | 1000001 |

Ensaio 2.14 Fatorial 3^4 - 1/3 repetição ($v = 27$; $b = 3$; $k = 9$):

Bloco Alternativas

- | | |
|-----|--|
| (1) | 0000 0122 0211 1022 1111 1200 2011 2100 2222 |
| (2) | 0021 0110 0202 1010 1102 1221 2002 2121 2210 |
| (3) | 0012 0101 0220 1001 1120 1212 2020 2112 2201 |

Ensaio 2.15 Fatorial 3^4 - 1/3 repetição ($v = 24$; $b = 6$; $k = 4$):

Bloco Alternativas

- | | | | |
|-----|---------------------|-----|---------------------|
| (1) | 0110 0101 1010 1001 | (4) | 2100 0021 0012 1200 |
| (2) | 0202 2020 2002 0220 | (5) | 0211 2011 1102 1120 |
| (3) | 1212 2112 2121 1221 | (6) | 2210 2201 0122 1022 |

3. Experimentos fatoriais confundidos em quadrados quasi-latino

Estes projetos se assemelham aos quadrados latinos ordinários uma vez que as diferenças entre as linhas e colunas dos quadrados se confundem e eliminam os erros experimentais dos efeitos de tratamentos confundidos; estes projetos não apresentam as propriedades típicas dos quadrados latino: que cada tratamento aparece uma vez em cada linha e cada coluna.

Ensaio 3.1 Fatorial 2^3 - ($v = 8$; $b = 2$; $k = 4$):

Bloco Alternativas

- | | |
|-----|-----------------|
| (1) | 000 101 011 110 |
| (2) | 100 010 001 111 |

Ensaio 3.2 Fatorial 2^4 - ($v = 16$; $b = 2$; $k = 8$):

Bloco Alternativas

- | | |
|-----|---|
| (1) | 0010 1101 0001 0111 1000 1110 0100 1011 |
| (2) | 1111 0000 0110 1001 0101 1010 1100 0011 |

Ensaio 3.3 Fatorial 2^5 - ($v = 32$; $b = 4$; $k = 8$):

Bloco Alternativas

- | | |
|-----|---|
| (1) | 00000 01101 00111 01010 11100 10110 11011 10100 |
| (2) | 11001 10100 11110 10011 00101 01111 00010 01000 |
| (3) | 01100 00001 01011 00110 10111 11101 10000 11010 |
| (4) | 10101 11000 10010 11111 01110 00100 01001 00011 |

4. Projetos de Blocos em Cadeia Generalizada (BCG)

Mandel² (1954) demonstrou que um projeto de blocos em cadeia podem escrever-se de tal modo que seja possível ajustar as diferenças entre da linhas e colunas.

O total de alternativa será formada com a rotação completa da atribuição dos atributos nas colunas correspondentes aos níveis e cada rotação se constitui em um bloco.

Ensaio 4.1 *Níveis = 3, Atributos = 5, Alternativas por blocos = 3, E = 0.96*

Alternativas	Atributos / Níveis	(Utiliza-se em experimento 3^3 , 3^4 e 3^5)
(1)	1 2 3 2 3	
(2)	2 3 1 1 2	
(3)	3 1 2 3 1	

Ensaio 4.2 *Níveis = 3, Atributos = 7, Alternativas por blocos = 3, E = 0.98*

Alternativas	Atributos / Níveis	(Utiliza-se em experimento 3^6 e 3^7)
(1)	1 2 3 1 2 3 1	
(2)	2 3 1 3 1 2 2	
(3)	3 1 2 2 3 1 3	

Ensaio 4.3 *Níveis = 4, Atributos = 5, Alternativas por blocos = 4, E = 0.98*

Alternativas	Atributos / Níveis	(Utiliza-se em experimento 4^3 , 4^4 e 4^5)
(1)	1 2 3 4 1	
(2)	2 3 4 1 4	
(3)	3 4 1 2 3	
(4)	4 1 2 3 2	

Ensaio 4.4 *Níveis = 4, Atributos = 7, Alternativas por bloco = 4, E = 0.98*

Alternativas	Atributos / Níveis	(Utiliza-se em experimento 4^6 e 4^7)
(1)	1 2 3 4 1 2 3	
(2)	2 1 4 3 3 4 1	
(3)	3 4 1 2 4 3 2	
(4)	4 3 2 1 2 1 4	

² MANDEL, J. Chain block designs with two-way elimination of heterogeneity. *Biometrics*. 10, 251-272, 1954.

Anexo 4

CONJUNTO DE ESCOLHA COM PROBABILIDADE CONDICIONAL

Os arranjos das alternativas em blocos, constante da tabela A4.1, foram montados com base nos arranjos desenvolvidos de acordo com as referências e adaptados para serem utilizados nas Técnicas de Preferência Declarada, com a divisão do conjunto das alternativas em blocos (conjunto de escolha com probabilidade condicional) para serem submetidos ao entrevistado, para o processo de escolha, ordenação ou avaliação.

Tabela A4.1 – Resumo dos Arranjos em Blocos

Nr	T	k	B	R	λ_1	λ_2	E	Projeto	Obs.
01	6	3	10	3	2	-	0,80	2.1	BIB
02	7	3	7	3	1	-	0,78	2.2	BIB
03		3	7	3	1	-	0,78	4.1	BIB
04		4	7	4	2	-	0,88	2.3	BIB
05		4	7	4	1	-	0,88	4.2	BIB
06	8	3	8	3	1	-	0,76	5.1	BIPB
07		4	6	3	2	0	0,86	5.2	BIPB – Geometria Finita
08		4	14	7	3	-	0,86	2.4	BIB em Repetição
09	9	3	9	3	1	-	0,75	5.3	BIPB
10		3	12	4	1	-	-	1.1	Reticulados quadrados
11		4	2x9	4	1	-	0,84	3.2	Grupos de repetição
12		4	2x9	8	3	-	0,84	2.5	Grupos de repetição
13		4	15	6	2	-	0,83	2.8	BIB
14		4	9	4	1	-	0,84	5.4	BIPB
15		5	2x9	10	5	-	0,90	2.6	Grupos de repetição
16		6	12	8	5	-	0,94	2.7	Grupos de repetição
17	10	4	10	4	1	0	0,83	5.5	BIPB (*)
18	11	5	11	5	2	-	0,88	2.9	BIB
19		5	11	5	2	-	0,88	4.3	BIB
20		6	11	6	3	-	0,92	3.1	Quadrado Latino Incompleto
21	12	3	3x4	3	1	-	-	1.3	Reticulado retangular
22		4	9	3	1	0	0,80	5.6	BIPB (*)
23	13	4	13	4	1	-	0,81	4.4	BIB
24		4	13	4	1	-	0,81	2.10	BIB
25	14	6	7	3	1	-	0,92	5.7	BIPB (*)
26	15	4	15	4	1	0	0,79	5.8	BIPB
27	16	4	12	3	1	-	0,80	5.9	BIPB (*)
28		4	20	5	1	-	-	1.2	Reticulado quadrado
29		6	2x12	9	3	-	0,89	2.11	Grupos de repetições
30		6	16	6	2	-	0,89	4.5	BIB
31	18	6	12	4	12	1	0,78	5.10	BIPB (*)
32	20	4	3x5	3	-	-	-	1.4	Reticulados retangular 4x5

(*) Arranjos adaptados pelo autor da tese.

1. Projetos de Experimentos Reticulados (“Lattice”)

Reticulados balanceados: O número de tratamento deverá ser um quadrado exato. O tamanho do bloco é a raiz quadrado. Todos pares de alternativas são comparados com o mesmo grau de precisão (COCHRAN e COX, 1978):

Projeto 1.1 Reticulados 3×3 , $t = 9$, $k = 3$, $r = 4$, $b = 12$, $\lambda = 1$;

Blocos

(1) 1 2 3	(4) 1 4 7	(7) 1 5 9	(10) 1 8 6
(2) 4 5 6	(5) 2 5 8	(8) 7 2 6	(11) 4 2 9
(3) 7 8 9	(6) 3 6 9	(9) 4 8 3	(12) 7 5 3

Projeto 1.2 Reticulados 4×4 , $t = 16$, $k = 4$, $r = 5$, $b = 20$, $\lambda = 1$;

Blocos

(1) 1 2 3 4	(6) 2 6 10 14	(11) 9 14 3 8	(16) 9 6 15 4
(2) 5 6 7 8	(7) 3 7 11 15	(12) 13 10 7 4	(17) 1 10 15 8
(3) 9 10 11 12	(8) 4 8 12 16	(13) 1 14 7 1	(18) 9 2 7 16
(4) 13 14 15 16	(9) 1 6 11 16	(14) 13 2 11 8	(19) 13 6 3 12
(5) 1 5 9 13	(10) 5 2 15 12	(15) 5 10 3 16	(20) 5 14 11 4

Reticulados retangulares: projetos desenvolvidos para $k(k+1)$ alternativas em blocos de k unidades. Os blocos poderão ser formados tanto tomando-se as linhas como as colunas (HARSHBARGER, 1949):

Projeto 1.3 Reticulados 3×4 , $t = 12$, $k = 3$, $r = 3$, $b = 3 \times 4$;

Rep. 1

(X1) 1 2 3
(X2) 4 5 6
(X3) 7 8 9
(X4) 10 11 12

Rep. 2

(Y1) 4 7 10
(Y2) 1 8 11
(Y3) 2 5 12
(Y4) 3 6 9

Rep. 3

(Z1) 6 8 12
(Z2) 2 9 10
(Z3) 3 4 11
(Z4) 1 5 7

Projeto 1.4 Reticulados 4×5 , $t = 20$, $k = 4$, $r = 3$, $b = 3 \times 5$;

Rep. 1

(X1) 1 2 3 4
(X2) 5 6 7 8
(X3) 9 10 11 12
(X4) 13 14 15 16
(X5) 17 18 19 20

Rep. 2

(Y1) 5 9 13 17
(Y2) 1 10 14 18
(Y3) 2 6 15 19
(Y4) 3 7 11 20
(Y5) 4 8 12 16

Rep. 3

(Z1) 8 11 15 18
(Z2) 2 9 16 20
(Z3) 4 7 14 17
(Z4) 1 5 12 19
(Z5) 3 6 10 13

2. Experimentos com alternativa grupada em Blocos Incompletos

Como seu nome indica, estes projetos introduzidos por YATES (1936), estão arranjados em blocos ou grupos menores que uma repetição completa e permite eliminar a heterocedasticidade em uma quantidade maior do que é possível com blocos ao acaso ou quadrados latinos.

Os projetos de Blocos Incompletos podem ser balanceados ou parcialmente balanceados. Os projetos balanceados se caracterizam pelo fato que cada par de alternativas ocorrem λ vez juntas nos blocos.

Os projetos de Blocos Incompletos estão classificados conforme o tipo:

- Tipo I – Projetos arranjados em repetições (utilizar pelo menos duas repetições);
- Tipo II – Projetos arranjados em grupos de repetições (utilizar pelo menos um dos grupos);
- Tipo III – Projetos não arranjados em repetições ou grupos de repetições (BIB);
- Tipo IV – Experimentos com $t = b$ (BIB);
- Tipo V – Experimentos pequenos.

Projeto 2.1 $t = 6, k = 3, r = 5, b = 10, \lambda = 2, E = 0.80, \text{Tipo III}$

Blocos Alternativas

(1) 1 2 5	(3) 1 3 4	(5) 1 4 5	(7) 2 3 5	(9) 3 5 6
(2) 1 2 6	(4) 1 3 6	(6) 2 3 4	(8) 2 4 6	(10) 4 5 6

Projeto 2.2 $t = 7, k = 3, r = 3, b = 7, \lambda = 1, E = 0.78, \text{Tipo V}$

Blocos Alternativas

(1) 1 2 4	(3) 3 4 6	(5) 5 6 1	(7) 7 1 3
(2) 2 3 5	(4) 4 5 7	(6) 6 7 2	

Projeto 2.3 $t = 7, k = 4, r = 4, b = 7, \lambda = 2, E = 0.88, \text{Tipo V}$

Blocos Alternativas

(1) 3 5 6 7	(3) 1 2 5 7	(5) 2 3 4 7	(7) 2 4 5 6
(2) 1 4 6 7	(4) 1 2 3 6	(6) 1 3 4 5	

Projeto 2.4 $t = 8, k = 4, r = 7, b = 14, \lambda = 3, E = 0.86, \text{Tipo I}$

Rep. 1

(1) 1 2 3 4
(2) 5 6 7 8

Rep. 2

(3) 1 2 7 8
(4) 3 4 5 6

Rep. 3

(5) 1 3 6 8
(6) 2 4 5 7

Rep. 4

(7) 1 4 6 7
(8) 2 3 5 8

Rep. 5

(9) 1 2 5 6

(10) 3 4 7 8

Rep. 6

(11) 1 3 5 7

(12) 2 4 6 8

Rep. 7

(13) 1 4 5 8

(14) 2 3 6 7

Projeto 2.5 $t = 9$, $k = 4$, $r = 8$, $b = 18$, $\lambda = 3$, $E = 0.84$, Tipo II

Rep. 1,2,3 e 4

(1) 1 4 6 7 (6) 4 5 6 9

(2) 2 6 8 9 (7) 2 3 6 7

(3) 1 3 8 9 (8) 2 4 5 8

(4) 1 2 3 4 (9) 3 5 7 9

(5) 1 5 7 8

Rep. 5,6,7 e 8

(10) 1 2 5 7

(11) 2 3 5 6

(12) 3 4 7 9

(13) 1 2 4 9

(14) 1 5 6 9

(15) 1 3 6 8

(16) 4 6 7 8

(17) 3 4 5 8

(18) 2 7 8 9

Projeto 2.6 $t = 9$, $k = 5$, $r = 10$, $b = 18$, $\lambda = 5$, $E = 0.90$, Tipo II

Rep. 1,2,3,4 e 5

(1) 1 2 3 7 8

(2) 1 2 4 6 8

(3) 2 3 5 8 9

(4) 2 3 4 6 9

(5) 1 3 4 5 7

(6) 2 4 5 6 7

(7) 1 3 6 7 9

(8) 1 4 5 8 9

(9) 5 6 7 8 9

Rep. 6,7,8,9 e 10

(10) 1 2 3 5 9

(11) 1 2 5 6 8

(12) 1 3 4 5 6

(13) 2 3 4 7 8

(14) 2 4 5 7 9

(15) 3 5 6 7 8

(16) 1 4 7 8 9

(17) 3 4 6 8 9

(18) 1 2 6 7 9

Projeto 2.7 $t = 9$, $k = 6$, $r = 8$, $b = 12$, $\lambda = 5$, $E = 0.94$, Tipo II

Rep. 1 e 2

(1) 1 2 4 5 7 8

(2) 2 3 5 6 8 9

(3) 1 3 4 6 7 9

Rep. 3 e 4

(4) 1 2 5 6 7 9

(5) 1 3 4 5 8 9

(6) 2 3 4 6 7 8

Rep. 5 e 6

(7) 1 3 5 6 7 8

(8) 1 2 4 6 8 9

(9) 2 3 4 5 7 9

Rep. 7 e 8

(10) 4 5 6 7 8 9

(11) 1 2 3 4 5 6

(12) 1 2 3 7 8 9

Projeto 2.8 $t = 9$, $k = 4$, $r = 6$, $b = 15$, $\lambda = 2$, $E = 0.83$, Tipo III

Blocos

(1) 1 2 3 4

(2) 1 2 5 6

(3) 1 3 7 8

(4) 1 4 9 10

(5) 1 5 7 9

(6) 1 6 8 10

(7) 2 3 6 9

(8) 2 4 7 10

(9) 2 5 8 10

(10) 2 7 8 9

(11) 3 5 9 10

(12) 3 6 7 10

(13) 3 4 5 8

(14) 4 5 6 7

(15) 4 6 8 9

Projeto 2.9 $t = 11$, $k = 5$, $r = 5$, $b = 11$, $\lambda = 2$, $E = 0.88$, Tipo IV

Blocos

(1) 1 2 3 5 8	(5) 5 6 7 9 1	(9) 2 5 9 10 11
(2) 2 3 4 6 9	(6) 2 6 7 8 10	(10) 1 3 6 10 11
(3) 3 4 5 7 10	(7) 3 7 8 9 11	(11) 1 2 4 7 11
(4) 4 5 6 8 11	(8) 1 4 8 9 10	

Projeto 2.10 $t = 13$, $k = 4$, $r = 4$, $b = 13$, $\lambda = 1$, $E = 0.81$, Tipo IV

Blocos

(1) 1 2 4 10	(5) 5 6 8 1	(9) 5 9 10 1	(13) 1 3 9 13
(2) 2 3 5 11	(6) 2 6 7 9	(10) 6 10 11 13	
(3) 3 4 6 12	(7) 3 7 8 10	(11) 1 7 11 12	
(4) 4 5 7 13	(8) 4 8 9 112	(12) 2 8 12 13	

Projeto 2.11 $t = 16$, $k = 6$, $r = 9$, $b = 24$, $\lambda = 3$, $E = 0.89$, Tipo II

Rep. 1,2 e 3

(1) 1 2 5 6 11 12
(2) 3 4 7 8 9 10
(3) 5 6 9 10 13 14
(4) 7 8 11 12 15 16
(5) 1 2 9 10 15 16
(6) 3 4 11 12 13 14
(7) 1 2 7 8 13 14
(8) 3 4 5 6 15 16

Rep. 4,5 e 6

(9) 1 3 6 8 13 15
(10) 2 4 5 7 14 15
(11) 5 7 9 11 13 15
(12) 6 8 10 12 14 16
(13) 2 4 6 8 9 11
(14) 1 3 5 7 10 12
(15) 2 4 10 12 13 15
(16) 1 3 9 11 14 16

Rep. 7,8 e 9

(17) 1 4 5 8 10 11
(18) 2 3 6 7 9 12
(19) 5 8 9 12 13 16
(20) 1 4 6 7 13 16
(21) 1 4 9 12 14 15
(22) 6 7 10 11 14 15
(23) 2 3 10 11 13 16
(24) 2 3 5 8 14 15

3. Quadros Latinos Incompletos

Estes projetos, que se constroem por um rearranjo de alguns blocos incompletos balanceado, possuem o “doble control” característico do quadrado latino sem a restrição de que o número de repetições tenha que ser igual ao número de tratamento (COCHRAN e COX, 1978):

Projeto 3.1 $t = 11$, $k = 6$, $r = 6$, $b = 11$, $\lambda = 3$, $E = 0.92$, Tipo I

Blocos

(1) 6 7 8 9 10 11
(2) 5 8 4 11 2 9
(3) 4 5 7 3 11 10
(4) 3 10 2 6 5 8

Blocos

(5) 2 3 9 7 4 6
(6) 1 6 10 4 9 5
(7) 9 1 3 5 8 7
(8) 8 2 1 10 7 4

(9) 7 11 5 1 6 2
(10) 11 4 6 8 1 3
(11) 10 9 11 2 3 1

Projeto 3.2 $t = 9$, $k = 4$, $r = 8$, $b = 18$, $\lambda = 3$, $E = 0.84$, Tipo V

Rep. 1,2,3 e 4

(1) 1 4 6 7

(2) 2 6 8 9

(3) 3 8 9 1

(4) 4 1 3 2

(5) 5 7 1 8

(6) 6 9 4 5

(7) 7 3 2 6

(8) 8 2 5 4

(9) 9 5 7 3

Rep. 5,6,7 e 8

(10) 1 2 5 7

(11) 2 3 6 5

(12) 3 4 7 9

(13) 4 9 2 1

(14) 5 1 9 6

(15) 6 8 1 3

(16) 7 6 4 8

(17) 8 5 3 4

(18) 9 7 8 2

4. Blocos Incompletos Balanceado desenvolvidos por Yates (BIB)

Projeto 4.1 $t = 7$, $k = 3$, $r = 3$, $b = 7$, $\lambda_1 = 1$, $E = 0.78$ Tipo BIB

Blocos

(1) 1 2 3

(3) 1 6 7

(5) 3 5 7

(7) 3 5 6

(2) 1 4 5

(4) 2 4 6

(6) 3 4 7

Projeto 4.2 $t = 7$, $k = 4$, $r = 4$, $b = 7$, $\lambda_1 = 1$, $\lambda_2 = 0$, Tipo BIB

Blocos

(1) 1 2 3 4

(3) 1 3 5 7

(5) 2 3 6 7

(7) 3 4 5 6

(2) 1 2 5 6

(4) 1 4 6 7

(6) 2 4 5 7

Projeto 4.3 $t = 11$, $k = 5$, $r = 5$, $b = 11$, $\lambda_1 = 2$, $E = 0.88$ Tipo BIB Yates

Blocos

(1) 1 2 3 4 5

(5) 1 5 8 10 11

(9) 3 4 6 8 11

(2) 1 2 6 7 8

(6) 2 3 7 10 11

(10) 3 5 7 8 9

(3) 1 3 6 9 10

(7) 2 4 8 9 10

(11) 4 5 6 7 10

(4) 1 4 7 9 11

(8) 2 5 6 9 11

Projeto 4.4 $t = 13$, $k = 4$, $r = 4$, $b = 13$, $\lambda_1 = 1$, $e = 0.81$ BIB

Blocos

(1) 1 2 3 4

(5) 2 5 8 11

(9) 3 6 10 11

(13) 4 7 9 11

(2) 1 5 6 7

(6) 2 6 9 12

(10) 3 7 8 12

(3) 1 8 9 10

(7) 2 7 10 13

(11) 4 5 10 12

(4) 1 11 12 13

(8) 3 5 9 13

(12) 4 6 8 13

Projeto 4.5 $t = 16, k = 6, r = 6, b = 16, \lambda_1 = 2, e = 0,89$ BIB

Blocos

(1) 1 2 3 4 5 6	(7) 2 3 7 14 15 16	(13) 3 6 8 9 13 14
(2) 1 2 7 8 9 10	(8) 2 4 8 12 13 16	(14) 4 5 7 10 13 14
(3) 1 3 7 11 12 19	(9) 2 5 9 11 13 15	(15) 4 6 7 9 12 15
(4) 1 4 8 11 14 15	(10) 2 6 10 11 12 14	(16) 5 6 7 8 11 16
(5) 1 5 9 12 14 16	(11) 3 4 9 10 11 16	
(6) 1 6 10 13 15 16	(12) 3 5 8 9 13 14	

5. Blocos Incompletos Parcialmente Balanceado (BIPB)

Obs. Arranjos desenvolvidos com base no Artigo de BOSE e NAIR (1939) e no livro de KEMPTHORNE (1967), úteis na divisão do conjunto de escolha em blocos:

Projeto 5.1 $t = 8, k = 3, r = 3, b = 8, \lambda_1 = 1, \lambda_2 = 0, E = 0,78$ Tipo BIPB

Blocos

(1) 1 2 4	(3) 1 3 5	(5) 8 7 3	(7) 2 3 6
(2) 1 7 6	(4) 8 2 5	(6) 8 4 6	(8) 7 4 5

Projeto 5.2 $t = 8, k = 4, r = 3, b = 6, \lambda_1 = 2, \lambda_2 = 0, E = 0,82$ BIPB

Blocos

(1) 1 2 3 4	(3) 1 4 5 8	(5) 3 4 7 8
(2) 1 2 5 6	(4) 2 3 6 7	(6) 5 6 7 8

Projeto 5.3 $t = 9, k = 3, r = 3, b = 9, \lambda_1 = 1, \lambda_2 = 0, E = 0,75$ Tipo BIPB

Blocos

(1) 1 2 3	(3) 1 7 5	(5) 6 9 5	(7) 7 9 3	(9) 2 9 4
(2) 1 6 4	(4) 6 8 3	(6) 7 8 4	(8) 2 8 5	

Projeto 5.4 $t = 9, k = 4, r = 4, b = 9, \lambda_1 = 1, \lambda_2 = 0, E = 0,84$ BIPB

Blocos

(1) 2 3 4 7	(4) 1 5 6 7	(7) 1 4 8 9
(2) 1 3 5 8	(5) 2 4 6 8	(8) 2 5 7 9
(3) 1 2 6 9	(6) 3 4 5 9	(9) 3 6 7 8

Projeto 5.5 $t = 10$, $k = 4$, $r = 4$, $b = 10$, $\lambda_1 = 1$, $\lambda_2 = 0$, $E = 0,83$ BIPB

Blocos

(1) 1 2 3 5	(4) 4 1 6 10	(7) 7 9 1 5	
(2) 2 4 8 6	(5) 5 8 7 3	(8) 8 10 2 7	(10) 10 6 5 4
(3) 3 7 4 9	(6) 6 9 2 8	(9) 9 3 10 1	

Projeto 5.6 $t = 12$, $k = 4$, $r = 3$, $b = 9$, $\lambda_1 = 1$, $\lambda_2 = 0$, $E = 0,82$ BIPB

Blocos

(1) 1 2 5 9	(4) 2 5 9 10	(7) 3 6 9 12
(2) 1 3 6 8	(5) 2 6 7 11	(8) 5 7 10 12
(3) 1 4 7 10	(6) 3 4 8 11	(9) 4 8 11 12

Projeto 5.7 $t = 14$, $k = 6$, $r = 3$, $b = 7$, $\lambda_1 = 1$, $\lambda_2 = 3$, $E = 0,92$ BIPB

Blocos

(1) 1 2 3 8 9 10	(4) 2 4 6 9 11 13	(7) 3 5 6 10 12 13
(2) 1 4 5 8 1 12	(5) 2 5 7 9 12 14	
(3) 1 6 7 8 13 14	(6) 3 4 7 10 11 14	

Projeto 5.8 $t = 15$, $k = 4$, $r = 4$, $b = 15$, $\lambda_1 = 1$, $\lambda_2 = 0$, $E = 0,79$ BIPB

Blocos

(1) 1 2 4 8	(6) 6 7 9 13	(11) 3 11 12 14
(2) 2 3 5 9	(7) 7 8 10 14	(12) 4 12 13 15
(3) 3 4 5 10	(8) 8 9 11 15	(13) 1 5 13 14
(4) 4 5 7 11	(9) 1 9 10 12	(14) 2 6 14 15
(5) 5 6 8 12	(10) 2 10 11 13	(15) 1 3 7 15

Projeto 5.9 $t = 16$, $k = 4$, $r = 3$, $b = 12$, $\lambda_1 = 1$, $\lambda_2 = 0$, $E = 0,80$ BIPB

Blocos

(1) 1 2 15 16	(4) 2 3 8 9	(7) 9 11 13 16	(10) 10 11 14 16
(2) 1 3 9 10	(5) 2 4 11 12	(8) 6 7 10 12	(11) 5 7 9 11
(3) 1 4 5 6	(6) 3 4 13 14	(9) 5 8 14 15	(12) 6 8 13 15

Projeto 5.10 $t = 18$, $k = 6$, $r = 4$, $b = 12$, $\lambda_1 = 1$, $\lambda_2 = 4$, $E = 0,78$ BIPB

Blocos

(1) 1 2 3 10 11 12	(5) 1 5 9 10 14 18	(9) 2 6 7 11 15 16
(2) 4 5 6 13 14 15	(6) 1 6 8 10 15 17	(10) 3 4 8 12 13 17
(3) 7 8 9 16 17 18	(7) 2 4 9 11 13 18	(11) 3 5 7 12 14 16
(4) 1 4 7 10 13 16	(8) 2 5 8 11 14 17	(12) 3 6 9 12 15 18

Anexo 5

EXEMPLOS DE UTILIZAÇÃO DOS ARRANJOS ORTOGONAIS

Exemplo 1: Delineamentos Experimental : 5 atributos com 2 níveis cada (2^5);

(os atributos podem ser discretos ou contínuos)

- 1) Utilizar a tabela A1 do anexo 2 e designar os 5 atributos às letras (colunas): A – B – C – D – E (número 13 do Anexo 1);

Conjunto de alternativas:

Alternativas	A	B	C	D	E
	β_1	β_2	β_3	β_4	β_5
01	0	0	0	0	0
02	0	0	1	1	0
03	0	1	0	1	1
04	0	1	1	0	1
05	1	0	0	1	1
06	1	0	1	0	1
07	1	1	0	0	0
08	1	1	1	1	0

- 2) Para realizar a pesquisa, pretende montar conjuntos de escolhas com 4 alternativas: utilizar o projeto 5.2 do anexo 4;

Conjunto de escolha: gera 6 conjuntos com 4 alternativas cada.

Projeto 6.2 $t = 8$, $k = 4$, $r = 3$, $b = 6$, $\lambda_1 = 2$, $\lambda_2 = 0$, $E = 0.82$ BIPB

Blocos

- | | | |
|-------------|-------------|-------------|
| (1) 1 2 3 4 | (3) 1 4 5 8 | (5) 3 4 7 8 |
| (2) 1 2 5 6 | (4) 2 3 6 7 | (6) 5 6 7 8 |

- Utilizar o Modelo LMPC;

O mesmo exemplo poderia ter sido efetuado com os arranjos fatoriais fracionários com divisão em blocos por confundimento através do ensaio 2.3 do anexo 3.

Exemplo 2: Delineamentos Experimental :

1 atributo com 4 níveis contínuo (A);

1 atributo com 3 níveis discretos (B);

3 atributos com 2 níveis discretos ou contínuos (C – D – E).

1) Utilizar a tabela A7 do anexo 2 (número 15 do anexo 1):

$$FU = \beta_1 X_1 + \beta_{21} X_{21} + \beta_{22} X_{22} + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

Tabela A7 – Ensaios fatoriais com atributos de 4 e 2 níveis: L16

Alternativas	A	B		C	D	E
	β_1	β_{21}	β_{22}	β_3	β_4	β_5
01	0	0	0	0	0	0
02	0	0	1	0	0	1
03	0	1	0	1	1	0
04	0	1	1	1	1	1
05	1	0	0	0	1	0
06	1	0	1	0	1	1
07	1	1	0	1	0	0
08	1	1	1	1	0	1
09	2	0	0	1	1	1
10	2	0	1	1	1	0
11	2	1	0	0	0	1
12	2	1	1	0	0	0
13	3	0	0	1	0	1
14	3	0	1	1	0	0
15	3	1	0	0	1	1
16	3	1	1	0	1	0

2) Formação do conjunto de escolha com 4 alternativas cada: Projeto 5.9 do anexo 4.

Projeto 5.9 $t = 16$, $k = 4$, $r = 3$, $b = 12$, $\lambda_1 = 1$, $\lambda_2 = 0$, $E = 0,80$ BIPB

Blocos

(1) 1 2 15 16

(4) 2 3 8 9

(7) 9 11 13 16

(10) 10 11 14 16

(2) 1 3 9 10

(5) 2 4 11 12

(8) 6 7 10 12

(11) 5 7 9 11

(3) 1 4 5 6

(6) 3 4 13 14

(9) 5 8 14 15

(12) 6 8 13 15

- Utilizar o Modelo LMPC.

Exemplo 3: Delineamentos Experimental :

1 atributo com 4 níveis contínuo (A);

4 atributos com 2 níveis discretos ou contínuos (B - C - D - E);

Existe um diferencial positivo ao passar do nível 0 para o nível 1 e assim por diante.

- 1) Utilizar a tabela A2 do anexo 2 (número 14 do anexo 1):

$$FU = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

Tabela A2 - Ensaios fatoriais com atributos de 4 e 2 níveis: L8

Alternativas	A	B	C	D	E
	β_1	β_2	β_3	β_4	β_5
01 (-)	0	0	0	0	0
02 (1)	0	1	1	1	1
03 (2)	1	0	1	0	1
04 (3)	1	1	0	1	0
05 (4)	2	0	1	1	0
06 (5)	2	1	0	0	1
07 (6)	3	0	0	1	1
08 (7)	3	1	1	0	0

Observação: Considerando-se que há uma vantagem quando os atributos passam do nível 0 para o nível 1, a alternativa 01 é totalmente dominada, em consequência, deve ser retirada da formação dos conjuntos de escolhas, no entanto deve fazer parte do delineamento por ocasião das estimativas dos parâmetros como sendo classificadas em último lugar.

- 2) Das 7 alternativas restantes, concorrerão para a formação dos conjuntos de escolhas com 4 alternativas cada: Projeto 2.3 do anexo 4.

Projeto 2.3 $t = 7$, $k = 4$, $r = 4$, $b = 7$, $\lambda = 2$, $E = 0.88$, Tipo V

Blocos Alternativas

(1) 3 5 6 7

(3) 1 2 5 7

(5) 2 3 4 7

(7) 2 4 5 6

(2) 1 4 6 7

(4) 1 2 3 6

(6) 1 3 4 5

- Utilizar o Modelo LMPC, incluir como última escolha a alternativa excluída e estimar com profundidade do rank até (k-2).

Anexo 6

ASPECTOS COMPUTACIONAIS DO MODELO LMPC

(Logit Multinomial com Probabilidade Condicional Explodida)

Passos para aplicar o Método do Algoritmo de Newton-Raphson.

Passo 0:

- atribuir um valor inicial aos $\beta_0 = \{\beta_{01}, \beta_{02}, \dots, \beta_{0k}\}$ (normalmente $\beta_0 = 0$);
- atribuir valores a e_1 e e_2 (normalmente $e_1 = 10^{-4}$ e $e_2 = 10^{-2}$).

Passo 1:

- linearize a função $\nabla L(\beta)$ em torno de β_w ;
- a aproximação da condição de primeira ordem é dada por

$$\nabla L(\beta_w) + \nabla^2 L(\beta_w) \cdot (\beta_{w+1} - \beta_w) = 0$$

Passo 2:

- Resolva a forma linearizada para

$$\beta_{w+1} = \beta_w - [\nabla^2 L(\beta_w)]^{-1} \cdot \nabla L(\beta_w).$$

Passo 3:

- Testar se $\beta_{w+1} - \beta_w$ é “pequeno”, ou seja:

$$\left[\sum_{k=1}^K (\beta_{w+1,k} - \beta_{w,k})^2 \right]^{1/2} \leq e_1 \text{ ou } \left| \frac{\beta_{w+1,k} - \beta_{w,k}}{\beta_{w,k}} \right| \leq e_2 \text{ para todo } k = 1, 2, \dots, K.$$

Se uma dessas condições é satisfeita, o processo termina com β_{w+1} sendo a solução. Em caso contrário fazer $w = w+1$ e voltar ao passo 1.

Seja a função Logit Multinomial com a probabilidade condicional:

$$L^* = \prod_{n=1}^N \frac{\sum_{i=1}^J e^{\beta \cdot \chi_{inb}}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta \cdot \chi_{inb}}} \times \prod_{j=1}^{J-1} \frac{e^{\beta \cdot \chi_{jnb}}}{\sum_{i \in C_b} e^{\beta \cdot \chi_{inb}}}$$

Aplicando o logaritmo na função L^* , temos a Função de Log-verossimilhança:

$$L = \sum_{n=1}^N \left[\ln \omega_b + \sum_{j=1}^{J-1} \left(\beta \chi_{jnb} - \ln \sum_{i=j}^J e^{\beta \chi_{inb}} \right) \right]$$

Calculando a primeira derivada de L em relação a um β_k :

$$\frac{\partial L}{\partial \beta_k} = [\nabla L]_k = 0$$

Sistema de Equações iguais a zero. $\frac{\partial L}{\partial \beta_k} = [\nabla L]_k = 0$

$$[\nabla L]_k = \sum_{n=1}^N \left[\frac{\sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbk}}{\sum_{i=1}^J e^{\beta' \chi_{inb}}} - \frac{\sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbk}}{\sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}}} + \sum_{j=1}^{J-1} \left(\chi_{jnbk} - \frac{\sum_{i=j}^J e^{\beta' \chi_{inb}} \chi_{inbk}}{\sum_{i=j}^J e^{\beta' \chi_{inb}}} \right) \right] = 0$$

E a Segunda derivada $\frac{\partial^2 L}{\partial \beta_k \partial \beta_l} = [\nabla^2 L]_{kl}$, onde

$$[\nabla^2 L]_{kl} = \sum_{n=1}^N \left[\frac{\sum_{i=1}^J e^{\beta' \chi_{inb}} \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbk} \chi_{inbl} - \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbk} \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbl}}{\sum_{i=1}^J e^{\beta' \chi_{inb}} \sum_{i=1}^J e^{\beta' \chi_{inb}}} - \right. \\ \left. + \frac{\sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}} \times \sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbk} \chi_{inbl} - \sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbk} \times \sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}} \chi_{inbl}}{\left(\sum_{b=1}^B \sum_{i=1}^J e^{\beta' \chi_{inb}} \right)^2} - \right. \\ \left. + \sum_{j=1}^{J-1} \frac{\sum_{i=j}^J e^{\beta' \chi_{inb}} \times \sum_{i=j}^J e^{\beta' \chi_{inb}} \chi_{inbk} \chi_{inbl} - \sum_{i=j}^J e^{\beta' \chi_{inb}} \chi_{inbk} \times \sum_{i=j}^J e^{\beta' \chi_{inb}} \chi_{inbl}}{\left(\sum_{i=j}^J e^{\beta' \chi_{inb}} \right)^2} \right]$$

$N = r.b$, onde: N = número total de entrevista; r = número de repetições de entrevistas com cada conjunto de cartões; b = número de conjunto de cartões.

Ponto de quase-cumeeira

A semelhança com a cumeeira de uma casa, vem melhor traduzir o problema em questão. Os efeitos aparentes são percebidos por ocasião das estimativas dos parâmetros, e provavelmente, estão relacionados com possíveis combinações lineares na matriz de delineamento do modelo.

Efeitos aparentes iniciais:

- a função utilidade converge para um ponto de máximo inicial;
- a matriz de variância converge para uma matriz mínima;
- o valor da estatística t tende para um ponto de máximo;
- os parâmetros tendem para um determinado valor;

A partir de uma certa iteração, normalmente a partir da terceira, e ainda, sem atingir os valores estipulados pela regra de parada, ocorre o seguinte:

- a função utilidade converge de forma assintótica suavemente para um máximo;
- a matriz de variância diverge para o infinito;
- os valores das estatísticas t dos parâmetros tendem a 0;
- os parâmetros, a cada iteração a partir dessa, sofrem um acréscimo uniforme de uma valor quase constante.

As sucessivas iterações provocam pequenas variações de maximização no valor da FU, refletida por acréscimos quase constante e uniformes nos parâmetros, o que caracteriza o ponto de quase-cumeeira para a FU.

Dada as características do processo, as regras de parada, descritas por BEN-AKIVA e LERMAN (1989), não são eficazes e há a necessidade de se incluir uma nova regra de parada, de forma que o processo seja interrompido na iteração i em que o valor da estatística t é máximo, ou seja, o ponto onde $\sum_k |t_{i,k}| > \sum_k |t_{i-1,k}|$.

PESQUISA DE OPINIÃO PÚBLICA

“Ao preencher este questionário
estarás contribuindo para o
desenvolvimento de uma nova técnica de
pesquisa, a qual incorpora parâmetros
de subjetividade na análise das
relações humanas e comerciais.”

* Não há necessidade da
identificação do entrevistado.

A presente pesquisa tem dois objetivos claramente definidos:

- Fazer uma aplicação didática das Técnicas de Preferência Declarada;
- Determinar as variáveis latentes: *PREÇO*, *QUALIDADE* e *CONVENIÊNCIA* do ponto de vista dos clientes de Supermercados.

Entendimento das variáveis latentes:

PREÇO	QUALIDADE	CONVENIÊNCIA
- Ofertas	- Higiene e limpeza	- Variedade dos produtos
- Promoções	- Qualidade dos produtos	- Estacionamento
- Sistema de crédito	- Qualidade do atendimento	- Distância / Proximidade
		- Lojas de conveniência

Parte I – Levantamento Sócio-Econômico

1) Sexo:

Masculino []	Feminino []
---------------	--------------

2) Qual sua escolaridade?

1º Grau []	2º Grau []	Superior []
-------------	-------------	--------------

3) Qual sua idade?

Até 20 anos []	De 21 a 35 anos []	De 36 a 50 anos []	Acima de 51 anos []
--------------------	------------------------	------------------------	-------------------------

4) Renda familiar (em Reais):

Até 600,00 []	De 601,00 a 1.200,00 []	De 1.201,00 a 2.400,00 []	Acima de 2.400,00 []
-------------------	-----------------------------	-------------------------------	--------------------------

5) Quantas pessoas têm em sua Família?

Até duas []	Três []	Quatro []	Cinco ou mais []
--------------	----------	------------	-------------------

Parte II - Pesquisa de Preferência Revelada

6) Quanto, aproximadamente, sua família gasta mensalmente com as compras em Supermercados? R\$ _____.

7) Qual o bairro onde você reside? _____.

8) Assinale dois Supermercados onde normalmente realiza suas compras mensais:

SUPERMERCADO		"X"
Extra Supermercado	SM-A1	
Condor	SM-A2	
Superbaratão	SM-B1	
Super Dal Pozzo	SM-B2	
Superpão	SM-A5	
Triunfante	SM-A6	
Dal Pozzo	SM-B3	
(*)		
(*)		

(*) Anotar outros Supermercados não incluídos na lista.

Parte II - Pesquisa de Preferência Declarada

VARIÁVEL PREÇO

- 9) Assinale com um “X” na classe de produtos onde você gostaria que houvesse uma redução de 20% no preço.

Produtos	“X”	Usar o conjunto de Cartões
01 – Carnes		<i>Verde</i>
02 – Frutas e Verduras		<i>Azul</i>
03 – Bens de Mercearia		<i>Amarelo</i>
04 – Laticínios e Frios		<i>Vermelho</i>

Na linha onde assinalou um “X”, há uma indicação da cor do conjunto de Cartões do **Experimento Fatorial Preço**. De posse deste conjunto de cartões, deverá colocar na ordem de acordo com a sua preferência e assinalar as letras no quadro correspondente.

Lembrete: Cada cartão representa uma situação em um supermercado.

Ordem de Preferência	Letras correspondentes
1ª Preferência	
2ª Preferência	
Último cartão	

VARIÁVEL QUALIDADE

- 10) De posse do Cartões referente ao **Experimento Fatorial Qualidade**, ordenar os cartões de acordo com suas preferências e assinalar no quadro abaixo as letras correspondentes:

Ordem de Preferência	Letras correspondentes
1ª Preferência	
2ª Preferência	
3ª Preferência	
4ª Preferência	
5ª Preferência	
Último cartão	

VARIÁVEL CONVENIÊNCIA

- 11) De posse dos Cartões referente ao **Experimento Fatorial Conveniência**, ordenar os cartões de acordo com suas preferências e assinalar no quadro abaixo as letras correspondentes:

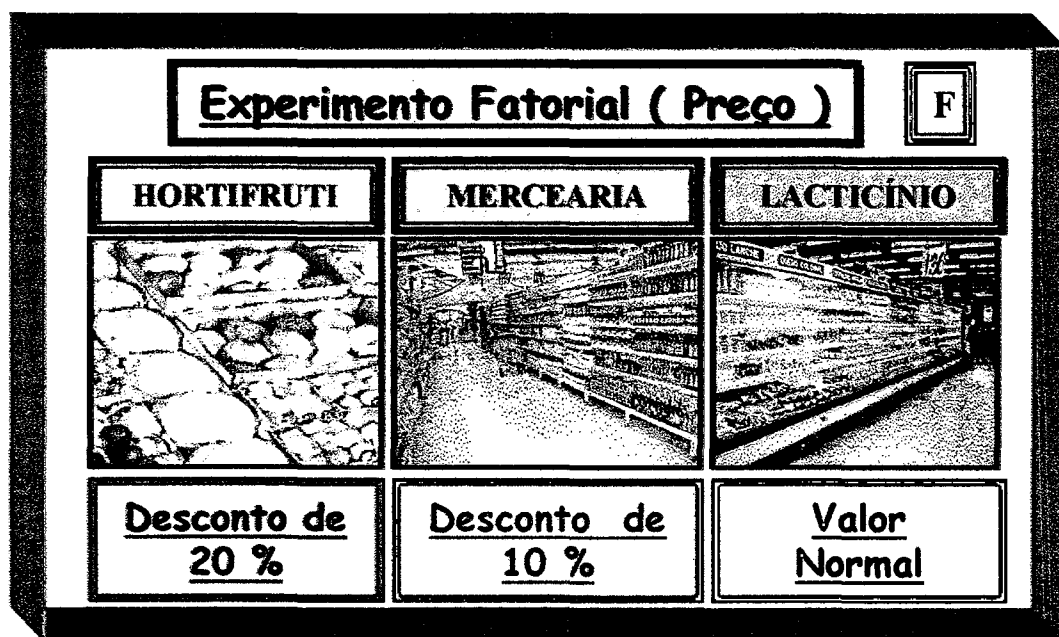
Conjunto n. ()

Ordem de Preferência	Letras correspondentes
1ª Preferência	
2ª Preferência	
3ª Preferência	
Último Cartão	

- 12) Assinale, respectivamente, 1, 2 e 3 de acordo com a ordem de importância dada aos atributos abaixo, quando se leva em conta a escolha de um Supermercado para fazer as compras mensais

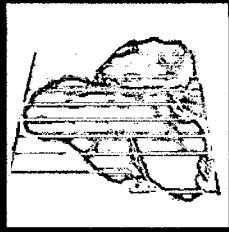
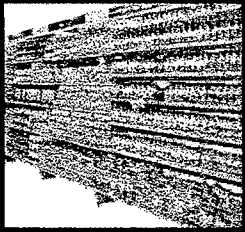
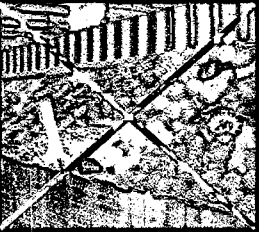



“O Pesquisador agradece a colaboração.”



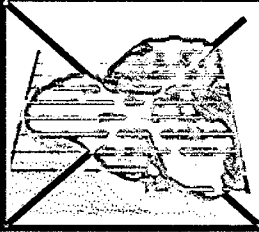
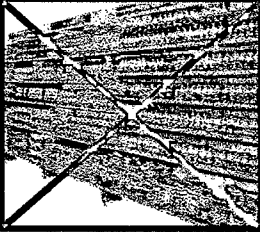
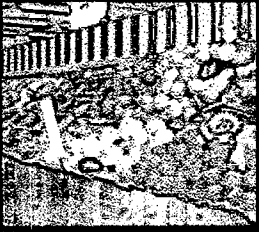

Experimento Fatorial (Qualidade)

S

CARNES	MERCEARIA	HORTIFRUTI	LACTICINIO
			
Cortes Variados Muitas Escolha Balcões Especiais	Produtos Variados Bem Organizado Estantes de Primeira	Produtos Básicos Pouco Organizado Estantes Simples	Produtos Basicos Qualidade Indefinida Balcões Simples

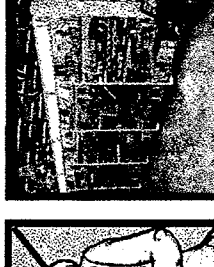
Experimento Fatorial (Qualidade)

K

CARNES	MERCEARIA	HORTIFRUTI	LACTICINIO
			
Cortes Limitados Pouca Escolha Balcões Simples	Produtos Básicos Pouco Organizado Estantes Simples	Produtos Variados Artigos Especiais Estantes Refrigeradas	Produtos Variados Primeira Qualidade Estantes Ideais

K

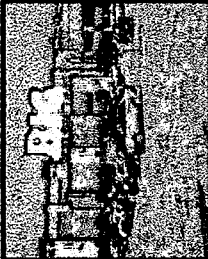

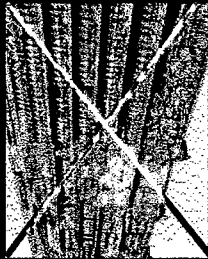



Diversificação



Vários tipos de Produtos

Os preços dos produtos são 5 % abaixo do valor normal.

M

Estacio- namento		Lojas de Conveniência		Organização		Sistema Crédito		Níveis de Serviço		Diversificação	
Amplo e com Segurança		Não tem		Apertado e Pouco Organizado	Com Vários Sistemas de Créditos	Promoções e Pessoal Disponíveis	Vários tipos de Produtos				

Os produtos são vendidos ao preço normal.

Anexo 9 – Relatório das Entrevistas

Atributos sociais: Sexo – Escol. – Idade – R. Familiar – Pessoas na família – Bairro – Despesas - PR

Nr	Atr. Sociais	Dsp	P.R.	Preço	Qualidade	Conveniência	C	P	Q
1	2 2 3 2 3 1	400	3 2	1 1 2 3 13	1 3 6 5 4 2 7	1	4 1 2 3 8	1 3	1 2
2	1 3 1 4 3 9	700	1 2	1 2 3 1 13	5 3 1 4 6 2 7	2	1 2 5 6 8	2 3	2 1
3	1 3 2 3 3 4	500	1 5	3 8 7 9 13	3 1 5 4 6 2 7	3	1 7 3 5 8	3 3	2 1
4	1 2 2 1 4 4	280	3 4	1 2 3 1 13	1 5 4 3 6 2 7	1	4 1 6 7 8	4 3	2 1
5	1 2 2 4 2 9	350	2 3	1 3 2 1 13	5 3 1 6 2 4 7	2	6 2 1 5 8	2 3	2 1
6	1 3 1 4 3 9	700	1 2	1 2 3 1 13	1 5 3 4 2 6 7	3	2 4 5 7 8	6 3	2 1
7	2 3 1 1 4 3	150	3 4	3 7 8 9 13	5 2 1 4 3 6 7	1	6 4 3 5 8	7 3	1 2
8	2 3 1 1 3 3	160	3 4	3 7 8 9 13	4 6 2 1 5 3 7	2	1 2 4 3 8	1 3	1 2
9	2 2 3 2 4 4	270	3 4	1 2 3 1 13	1 5 3 6 2 4 7	3	6 5 1 2 8	2 1	3 2
10	1 2 1 2 4 10	450	4 3	1 3 1 2 13	3 5 1 2 6 4 7	1	7 5 3 1 8	3 3	2 1
11	1 3 1 2 4 10	400	4 3	1 3 1 2 13	1 6 4 5 3 2 7	2	7 6 4 1 8	4 3	2 1
12	1 2 2 4 2 9	350	2 3	1 3 2 1 13	5 1 3 2 4 6 7	3	6 2 7 3 8	5 3	2 1
13	2 3 1 3 4 1	300	1 5	1 2 3 1 13	3 5 1 2 4 6 7	1	5 2 7 4 8	6 2	3 1
14	1 2 1 3 3 8	280	3 2	4 11 10 12 13	3 5 1 2 4 6 7	2	3 5 4 6 8	7 3	1 2
15	2 1 3 1 3 8	200	3 4	2 5 4 6 13	3 2 5 6 1 4 7	3	2 4 1 3 8	1 1	3 2
16	2 1 1 1 4 8	200	3 9	2 5 4 6 13	3 2 5 1 6 4 7	1	2 1 5 6 8	2 3	2 1
17	2 3 1 2 3 1	250	3 2	1 2 3 1 13	5 6 3 1 2 4 7	2	7 5 3 1 8	3 3	1 2
18	1 3 2 1 1 8	300	8 2	3 9 8 7 13	1 5 3 2 4 6 7	3	4 1 7 6 8	4 3	2 1
19	2 3 1 3 4 11	300	3 1	1 2 3 1 13	1 5 2 3 4 6 7	1	3 2 7 6 8	5 3	2 1
20	1 3 2 2 2 2	250	3 2	3 7 9 8 13	5 3 1 2 4 6 7	2	7 4 5 2 8	6 2	1 3
21	2 3 1 1 4 10	300	3 9	2 5 4 6 13	1 5 3 2 4 6 7	3	4 6 3 5 8	7 3	2 1
22	1 3 2 1 2 2	200	3 4	1 3 1 2 13	3 1 5 2 6 4 7	1	1 3 2 4 8	1 3	1 2
23	2 1 2 2 1 3	100	3 1	3 7 8 9 13	3 1 5 2 6 4 7	2	5 1 2 6 8	2 3	1 2
24	2 3 1 2 3 10	280	3 2	4 11 10 12 13	3 2 5 6 1 4 7	3	1 5 3 7 8	3 1	3 2
25	2 2 1 3 4 2	300	3 5	1 3 2 1 13	3 1 5 4 2 6 7	1	7 4 6 1 8	4 3	2 1
26	2 1 1 3 4 2	250	3 6	1 3 2 1 13	3 5 1 2 4 6 7	2	6 2 7 3 8	5 3	2 1
27	1 3 1 3 3 4	450	5 3	4 9 7 8 13	3 6 2 1 5 4 7	3	2 4 5 7 8	6 3	1 2
28	2 2 1 3 4 1	300	1 5	1 2 3 1 13	5 3 1 2 4 6 7	1	3 5 4 6 8	7 2	3 1
29	2 2 2 1 2 10	150	3 1	1 1 3 2 13	3 5 1 2 6 4 7	2	2 4 1 3 8	1 3	2 1
30	1 3 1 3 3 4	450	5 3	4 9 7 8 13	3 6 2 5 1 4 7	3	2 5 6 1 8	2 3	1 2
31	2 1 2 2 1 3	100	3 1	3 7 8 9 13	3 1 5 4 6 2 7	1	5 1 3 7 8	3 3	1 2
32	2 2 2 1 2 10	150	3 1	1 1 3 2 13	3 5 1 6 2 4 7	2	7 4 1 6 8	4 3	2 1
33	2 2 1 3 3 2	600	3 6	1 1 3 2 13	1 5 3 6 4 2 7	3	7 6 2 3 8	5 3	2 1
34	2 3 2 3 3 8	250	4 2	3 7 9 8 13	4 2 1 3 5 6 7	1	4 2 7 5 8	6 3	2 1
35	2 1 4 1 3 2	200	4 3	2 1 3 2 13	3 2 5 1 4 6 7	2	5 4 3 6 8	7 1	3 2
36	2 2 3 2 3 2	600	3 1	1 1 2 3 13	1 5 2 3 6 4 7	3	4 1 2 3 8	1 3	1 2
37	2 1 2 3 3 8	250	4 2	3 7 9 8 13	4 2 5 1 3 6 7	1	2 5 1 6 8	2 3	2 1
38	1 3 1 2 3 1	250	3 2	1 2 3 1 13	5 1 6 3 4 2 7	2	5 1 7 3 8	3 3	1 2
39	2 3 2 1 2 2	200	3 4	1 3 1 2 13	1 5 3 6 4 2 7	3	7 4 6 1 8	4 3	1 2
40	2 1 2 2 4 1	230	5 3	1 3 2 1 13	3 1 5 2 6 4 7	1	2 3 7 6 8	5 3	2 1
41	2 1 3 1 4 2	250	3 2	3 7 9 8 13	3 2 5 6 1 4 7	2	2 5 4 7 8	6 2	1 3
42	2 3 1 2 4 2	400	3 2	1 2 1 3 13	3 2 5 6 1 4 7	3	4 6 3 5 8	7 3	2 1
43	1 3 1 3 4 11	400	2 1	1 3 2 1 13	3 5 4 2 1 6 7	1	1 2 3 4 8	1 3	2 1
44	2 3 1 3 4 11	300	3 1	1 2 3 1 13	3 5 1 6 2 4 7	2	2 5 6 1 8	2 3	2 1
45	1 2 2 3 4 2	500	3 5	4 12 11 10 13	5 1 3 2 4 6 7	3	5 1 7 3 8	3 3	2 1
46	2 3 1 2 4 2	400	3 2	1 2 1 3 13	5 1 2 3 6 4 7	1	7 4 6 1 8	4 3	2 1
47	2 3 2 4 4 8	400	3 1	1 2 1 3 13	3 5 1 2 6 4 7	2	3 6 7 2 8	5 3	2 1
48	2 1 3 2 3 5	350	3 2	2 4 6 5 13	5 1 3 2 4 6 7	3	7 4 5 2 8	6 3	2 1
49	2 1 2 1 1 8	300	3 7	1 3 1 2 13	1 5 3 2 6 4 7	1	3 5 4 6 8	7 3	1 2
50	2 1 3 2 3 5	350	3 2	2 4 6 5 13	5 6 3 4 5 2 7	2	1 2 3 4 8	1 3	2 1
51	1 2 2 3 4 2	500	3 5	4 12 11 10 13	5 1 3 2 4 6 7	3	5 6 2 1 8	2 3	2 1
52	2 3 1 4 4 7	600	2 5	4 12 11 10 13	5 3 1 2 6 4 7	1	7 5 3 1 8	3 3	2 1
53	2 1 2 4 4 7	200	4 3	1 3 2 1 13	3 2 5 6 1 4 7	2	1 4 7 6 8	4 1	3 2
54	1 2 2 2 3 1	300	2 5	3 7 8 9 13	1 5 3 6 4 2 7	3	3 6 7 2 8	5 2	1 3
55	2 3 1 4 4 7	600	2 5	4 12 11 10 13	3 5 1 2 6 4 7	1	7 4 5 2 8	6 3	2 1
56	2 1 2 2 1 1	300	5 1	1 3 2 1 13	3 1 6 5 2 4 7	2	4 6 3 5 8	7 2	3 1
57	2 3 2 2 1 2	300	1 3	1 3 2 1 13	1 5 3 6 2 4 7	3	1 2 3 4 8	1 2	3 1
58	2 3 2 4 4 8	400	3 2	1 2 1 3 13	3 5 1 2 6 4 7	1	5 6 1 2 8	2 3	2 1
59	2 1 2 1 1 8	300	3 7	1 3 1 2 13	3 1 6 5 4 2 7	2	3 7 1 5 8	3 3	1 2
60	2 2 2 2 2 1	300	5 1	1 3 2 1 13	1 3 5 6 2 4 7	3	4 1 7 6 8	4 2	3 1
61	2 3 2 2 1 8	250	2 3	2 5 6 4 13	3 5 2 1 4 6 7	1	7 3 6 2 8	5 3	1 2
62	2 1 3 1 3 7	200	4 3	1 3 2 1 13	5 3 1 2 4 6 7	2	5 7 4 2 8	6 3	1 2
63	1 2 2 2 3 2	400	2 3	1 3 2 1 13	3 1 5 4 6 2 7	3	3 5 4 6 8	7 3	1 2
64	1 2 2 3 3 1	600	1 5	1 2 1 3 13	1 5 3 4 2 6 7	1	4 3 1 2 8	1 3	2 1

Nr	Atr. Sociais	Dsp	P.R.	Preço	Qualidade	Conveniência	C	P	Q
65	2 1 2 2 1 8	250	3 2	2 5 6 4 13	3 1 6 5 2 4 7	2	5 6 1 2 8	2 3	1 2
66	2 3 2 1 1 4	350	3 4	1 3 1 2 13	1 3 5 2 4 6 7	3	7 5 3 1 8	3 3	1 2
67	2 2 2 2 1 2	300	3 2	2 5 6 4 13	3 5 2 1 4 6 7	1	7 1 4 6 8	4 3	1 2
68	1 3 2 4 4 1	800	1 5	1 2 1 3 13	5 3 1 2 4 6 7	2	7 6 3 2 8	5 3	2 1
69	2 2 2 2 2 1	350	5 1	1 3 2 1 13	1 3 5 6 4 2 7	3	4 7 2 5 8	6 2	3 1
70	2 3 2 4 4 1	800	1 5	1 2 1 3 13	1 3 2 5 6 4 7	1	6 5 4 3 8	7 3	1 2
71	1 3 2 3 4 1	400	5 1	1 2 1 3 13	3 5 1 2 6 4 7	2	3 1 4 2 8	1 2	3 1
72	1 2 2 3 4 1	350	1 5	1 2 1 3 13	5 3 1 2 4 6 7	3	6 1 5 2 8	2 2	3 1
73	2 3 2 3 4 1	450	5 3	1 2 1 3 13	3 5 1 2 6 4 7	1	7 3 1 5 8	3 3	1 2
74	2 2 4 2 3 1	400	4 3	1 3 1 2 13	1 3 5 4 2 6 7	2	7 6 4 1 8	4 3	2 1
75	1 2 2 4 2 9	350	2 3	1 3 2 1 13	5 3 1 2 4 6 7	3	6 2 7 3 8	5 3	2 1
76	2 3 1 3 4 1	300	1 5	1 2 3 1 13	3 5 1 2 6 4 7	1	5 2 7 4 8	6 2	3 1
77	1 2 1 3 3 10	280	3 2	4 11 10 12 13	3 1 5 4 2 6 7	2	3 5 4 6 8	7 3	1 2
78	1 3 2 1 1 8	300	8 2	3 9 8 7 13	1 3 5 2 4 6 7	3	3 4 1 2 8	1 3	2 1
79	1 3 2 3 4 1	400	5 1	1 2 1 3 13	1 5 2 3 6 4 7	1	3 1 4 2 8	2 2	3 1
80	2 2 2 3 4 1	450	5 3	1 2 1 3 13	3 5 1 6 2 4 7	2	7 3 1 5 8	3 3	1 2
82	2 2 2 2 2 1	300	5 1	1 3 2 1 13	1 3 5 6 4 2 7	3	4 1 7 6 8	4 2	3 1
82	2 3 2 2 1 8	250	2 3	2 5 6 4 13	3 5 2 1 6 4 7	1	7 3 6 2 8	5 3	1 2
83	1 3 2 4 3 7	600	4 2	1 3 2 1 13	5 1 3 4 2 6 7	2	7 4 2 5 8	6 1	3 2
84	1 2 2 2 3 2	400	2 3	1 3 2 1 13	1 3 5 6 4 2 7	3	3 5 4 6 8	7 3	1 2
85	1 2 2 3 3 1	600	1 5	1 2 1 3 13	1 5 3 4 6 2 7	1	4 3 1 2 8	1 3	2 1
86	2 3 2 3 4 1	450	5 3	1 2 1 3 13	3 1 5 2 6 4 7	2	1 5 6 2 8	2 3	1 2
87	1 3 1 2 4 10	400	4 3	1 3 1 2 13	3 2 6 5 1 4 7	3	1 5 3 7 8	3 2	3 1
88	1 2 2 4 2 9	350	2 3	1 3 2 1 13	3 5 1 2 6 4 7	1	1 4 7 6 8	4 3	1 2
89	2 3 1 3 4 1	300	1 5	1 2 3 1 13	3 2 6 5 1 4 7	2	2 3 7 6 8	5 3	2 1
90	2 2 2 1 1 2	300	3 7	1 3 1 2 13	1 5 3 6 4 2 7	3	7 2 4 5 8	6 3	1 2
91	1 3 2 3 3 4	500	1 5	3 8 7 9 13	5 2 3 1 6 4 7	1	5 4 3 6 8	7 3	2 1
92	1 3 1 4 3 9	1000	1 2	1 2 3 1 13	3 5 1 2 4 6 7	2	2 4 3 1 8	1 3	2 1
93	1 3 2 3 3 4	500	1 5	3 8 7 9 13	3 2 5 6 4 1 7	3	1 2 6 5 8	2 3	2 1
94	1 2 2 1 4 4	280	3 4	1 2 3 1 13	3 5 2 1 6 4 7	1	1 5 3 7 8	3 3	2 1
95	1 2 2 4 2 9	350	2 3	1 3 2 1 13	3 5 1 6 2 4 7	2	1 4 6 7 8	4 3	2 1
96	1 3 1 4 3 9	1000	1 2	1 2 3 1 13	3 2 5 6 1 4 7	3	2 7 6 3 8	5 3	2 1
97	2 3 1 1 4 3	150	3 4	3 7 8 9 13	3 5 2 6 1 4 7	1	2 5 4 7 8	6 3	1 2
98	2 3 1 1 3 3	160	3 4	3 7 8 9 13	3 5 2 1 6 4 7	2	5 4 3 6 8	7 3	1 2
99	1 2 2 1 4 4	270	3 4	1 2 3 1 13	3 2 6 1 5 4 7	3	2 4 3 1 8	1 3	2 1
100	1 2 1 2 4 10	450	4 3	1 3 1 2 13	5 2 4 3 1 6 7	1	1 2 6 5 8	2 3	2 1
101	1 3 1 2 4 10	400	4 3	1 3 1 2 13	3 5 2 1 6 4 7	2	1 5 7 3 8	3 3	2 1
102	1 2 2 4 2 9	350	2 3	1 3 2 1 13	5 3 2 6 1 4 7	3	4 1 7 6 8	4 3	2 1
103	2 3 1 3 4 1	300	1 5	1 2 3 1 13	3 5 2 4 1 6 7	1	7 6 2 3 8	5 2	3 1
104	1 2 1 3 3 10	280	3 2	4 11 10 12 13	3 1 6 5 2 4 7	2	2 5 4 7 8	6 3	1 2
105	1 3 2 1 1 8	300	8 2	3 9 8 7 13	5 3 2 6 1 4 7	3	4 5 3 6 8	7 3	2 1
106	1 3 2 3 3 4	500	1 5	3 8 7 9 13	1 3 5 4 6 2 7	1	1 2 3 4 8	1 3	2 1
107	1 3 1 4 3 9	1000	1 2	1 2 3 1 13	5 3 1 4 6 2 7	2	1 2 5 6 8	2 3	2 1
108	1 3 2 3 3 4	500	1 5	3 8 7 9 13	3 1 5 4 6 2 7	3	1 7 3 5 8	3 3	2 1
109	1 2 2 1 4 4	280	3 4	1 2 3 1 13	1 5 4 3 6 2 7	1	4 1 6 7 8	4 3	2 1
110	1 2 2 4 2 9	350	2 3	1 3 2 1 13	5 3 1 6 2 4 7	2	6 2 1 5 8	2 3	2 1
111	1 3 1 4 3 9	1000	1 2	1 2 3 1 13	1 5 3 4 2 6 7	3	2 4 5 7 8	6 3	2 1
112	2 3 1 1 4 3	150	3 4	3 7 8 9 13	5 2 1 4 3 6 7	1	6 4 3 5 8	7 3	1 2
113	2 3 1 1 3 3	160	3 4	3 7 8 9 13	4 6 2 1 5 3 7	2	1 2 4 3 8	1 3	1 2
114	1 2 2 1 4 4	270	3 4	1 2 3 1 13	1 5 3 6 2 4 7	3	1 4 2 3 8	2 3	2 1
115	1 2 1 2 4 10	450	4 3	1 3 1 2 13	3 5 1 2 6 4 7	1	7 5 3 1 8	3 3	2 1
116	1 3 1 2 4 10	400	4 3	1 3 1 2 13	1 6 4 5 3 2 7	2	7 6 4 1 8	4 3	2 1
117	1 2 2 4 2 9	350	2 3	1 3 2 1 13	5 1 3 2 4 6 7	3	6 2 7 3 8	5 3	2 1
118	2 3 1 3 4 1	300	1 5	1 2 3 1 13	3 5 1 2 4 6 7	1	5 2 7 4 8	6 2	3 1
119	1 2 1 3 3 10	280	3 2	4 11 10 12 13	3 5 1 2 4 6 7	2	3 5 4 6 8	7 3	1 2
120	1 3 2 1 1 8	300	8 2	3 9 8 7 13	1 5 3 2 4 6 7	3	3 4 1 2 8	1 3	2 1
121	2 3 1 1 4 10	200	3 9	2 5 4 6 13	5 1 2 3 6 4 7	1	5 1 2 6 8	2 3	2 1
122	1 3 1 2 3 1	250	3 2	1 2 3 1 13	5 6 3 1 2 4 7	2	5 1 7 3 8	3 3	1 2
123	1 3 2 1 1 8	300	8 2	3 9 8 7 13	1 5 3 2 4 6 7	3	4 1 7 6 8	4 3	2 1
124	2 3 1 3 4 11	300	3 1	1 2 3 1 13	1 5 2 3 4 6 7	1	3 2 7 6 8	5 3	2 1
125	1 3 2 2 2 2	250	3 2	3 7 9 8 13	5 3 1 2 4 6 7	2	7 4 5 2 8	6 2	1 3
126	2 3 1 1 4 10	300	3 9	2 5 4 6 13	1 5 3 2 4 6 7	3	4 6 3 5 8	7 3	2 1
127	1 3 2 1 2 2	200	3 4	1 3 1 2 13	3 1 5 2 6 4 7	1	1 3 2 4 8	1 3	1 2
128	2 1 2 2 1 3	100	3 1	3 7 8 9 13	3 1 5 2 6 4 7	2	5 1 2 6 8	2 3	1 2
129	1 3 1 2 3 10	280	3 2	4 11 10 12 13	1 5 3 2 6 4 7	3	3 1 7 5 8	3 3	1 2
130	2 2 1 3 4 2	300	3 5	1 3 2 1 13	3 1 5 4 2 6 7	1	7 4 6 1 8	4 3	2 1
131	2 1 1 3 4 2	250	3 6	1 3 2 1 13	3 5 1 2 4 6 7	2	6 2 7 3 8	5 3	2 1
132	1 3 1 3 3 4	450	5 3	4 9 7 8 13	3 6 2 1 5 4 7	3	2 4 5 7 8	6 3	1 2
133	2 2 1 3 4 1	300	1 5	1 2 3 1 13	5 3 1 2 4 6 7	1	3 5 4 6 8	7 2	3 1
134	2 2 2 1 2 10	150	3 1	1 1 3 2 13	3 5 1 2 6 4 7	2	2 4 1 3 8	1 3	2 1

Nr	Atr. Sociais	Dsp	P.R.	Preço	Qualidade	Conveniência	C	P	Q
135	1 3 1 3 3 4	450	5 3	4 9 7 8 13	3 6 25 1 4 7	3	2 5 6 1 8	2 3	1 2
136	2 1 2 2 1 3	100	3 1	3 7 8 9 13	3 15 46 2 7	1	5 1 3 7 8	3 3	1 2
137	2 2 2 1 2 10	150	3 1	1 1 3 2 13	3 5 1 62 4 7	2	7 4 1 6 8	4 3	2 1
138	1 2 1 3 3 2	600	3 6	1 1 3 2 13	1 5 3 64 2 7	3	6 2 7 3 8	5 3	2 1
139	2 3 2 3 3 8	250	4 2	3 7 9 8 13	4 2 1 35 6 7	1	4 2 7 5 8	6 3	2 1
140	1 2 1 3 3 2	600	3 6	1 1 3 2 13	3 5 1 42 6 7	2	4 5 3 6 8	7 3	2 1
141	2 2 3 2 3 2	600	3 1	1 1 2 3 13	15 2 3 6 4 7	3	4 1 2 3 8	1 3	1 2
142	2 1 2 3 3 8	250	4 2	3 7 9 8 13	4 25 1 36 7	1	2 5 1 6 8	2 3	2 1
143	1 3 1 2 3 1	250	3 2	1 2 3 1 13	5 16 3 42 7	2	5 1 7 3 8	3 3	1 2
144	1 3 2 1 2 2	200	3 4	1 3 1 2 13	1 5 3 6 4 2 7	3	7 4 6 1 8	4 3	1 2
145	1 3 2 2 1 1	230	5 3	1 3 2 1 13	3 1 5 2 6 4 7	1	3 6 2 7 8	5 3	2 1
146	1 3 2 2 2 2	250	3 2	3 7 9 8 13	5 3 1 4 2 6 7	2	7 4 5 2 8	6 2	1 3
147	2 3 1 2 4 2	400	3 2	1 2 1 3 13	3 2 5 6 1 4 7	3	4 6 3 5 8	7 3	2 1
148	1 3 1 3 4 11	400	2 1	1 3 2 1 13	3 5 4 2 16 7	1	1 2 3 4 8	1 3	2 1
149	2 3 1 3 4 11	300	3 1	1 2 3 1 13	3 5 1 6 2 4 7	2	2 5 6 1 8	2 3	2 1
150	1 2 2 3 4 2	500	3 5	4 12 11 10 13	5 1 3 2 4 6 7	3	5 1 7 3 8	3 3	2 1
151	2 3 1 2 4 2	400	3 2	1 2 1 3 13	5 1 2 3 6 4 7	1	7 4 6 1 8	4 3	2 1
152	2 3 2 4 4 8	400	3 1	1 2 1 3 13	3 5 1 2 6 4 7	2	3 6 7 2 8	5 3	2 1
153	2 1 3 2 3 5	350	3 2	2 4 6 5 13	5 13 2 4 6 7	3	7 4 5 2 8	6 3	2 1
154	2 1 2 1 1 8	300	3 7	1 3 1 2 13	1 5 3 2 6 4 7	1	3 5 4 6 8	7 3	1 2
155	2 1 3 2 3 5	350	3 2	2 4 6 5 13	5 6 3 4 5 2 7	2	1 2 3 4 8	1 3	2 1
156	1 2 2 3 4 2	500	3 5	4 12 11 10 13	5 1 3 2 4 6 7	3	5 6 2 1 8	2 3	2 1
157	2 3 1 4 4 7	600	2 5	4 12 11 10 13	5 3 1 2 6 4 7	1	7 5 3 1 8	3 3	2 1
158	1 3 2 4 3 7	600	4 2	1 3 2 1 13	5 3 1 4 2 6 7	2	7 4 6 1 8	4 1	3 2
159	1 2 2 2 3 1	300	2 5	3 7 8 9 13	1 5 3 6 4 2 7	3	3 6 7 2 8	5 2	1 3
160	2 3 1 4 4 7	600	2 5	4 12 11 10 13	3 5 1 2 6 4 7	1	7 4 5 2 8	6 3	2 1
161	2 1 2 2 1 1	300	5 1	1 3 2 1 13	3 1 6 5 2 4 7	2	4 6 3 5 8	7 2	3 1
162	2 3 2 2 1 2	300	1 3	1 3 2 1 13	1 5 3 6 2 4 7	3	1 2 3 4 8	1 2	3 1
163	2 3 2 4 4 8	400	3 2	1 2 1 3 13	3 5 1 2 6 4 7	1	5 6 1 2 8	2 3	2 1
164	2 1 2 1 1 8	300	3 7	1 3 1 2 13	3 1 6 5 4 2 7	2	3 7 1 5 8	3 3	1 2
165	2 2 2 2 2 1	300	5 1	1 3 2 1 13	1 3 5 6 2 4 7	3	4 1 7 6 8	4 2	3 1
166	2 3 2 2 1 8	250	2 3	2 5 6 4 13	3 5 2 1 4 6 7	1	7 3 6 2 8	5 3	1 2
167	1 3 2 4 3 7	600	4 2	1 3 2 1 13	5 3 1 2 4 6 7	2	7 4 2 5 8	6 1	3 2
168	1 2 2 2 3 2	400	2 3	1 3 2 1 13	3 1 5 4 6 2 7	3	3 5 4 6 8	7 3	1 2
169	1 2 2 3 3 1	600	1 5	1 2 1 3 13	1 5 3 4 2 6 7	1	4 3 1 2 8	1 3	2 1
170	2 1 2 2 1 8	250	3 2	2 5 6 4 13	3 1 6 5 2 4 7	2	5 6 1 2 8	2 3	1 2
171	2 3 2 1 1 4	350	3 4	1 3 1 2 13	1 3 5 2 4 6 7	3	7 5 3 1 8	3 3	1 2
172	2 2 2 2 1 2	300	3 2	2 5 6 4 13	3 5 2 1 4 6 7	1	7 1 4 6 8	4 3	1 2
173	1 3 2 4 4 1	800	1 5	1 2 1 3 13	5 3 1 2 4 6 7	2	7 6 3 2 8	5 3	2 1
174	2 2 2 2 2 1	350	5 1	1 3 2 1 13	1 3 5 6 4 2 7	3	4 7 2 5 8	6 2	3 1
175	1 3 2 4 4 1	800	1 5	1 2 1 3 13	3 5 1 2 6 4 7	1	5 6 4 3 8	7 3	2 1

Dados Amostrais

Atributos Sociais	Níveis	Amostra
Sexo	Masculino	80 (46%)
	Feminino	95 (54%)
Escolaridade	1º Grau	27 (15%)
	2º Grau	58 (33%)
	3º Grau	90 (52%)
Idade	Até 20 anos	65 (37%)
	De 21 a 35 a.	97 (56%)
	De 26 a 50 a.	11 (06%)
	Acima de 50 a.	2 (01%)
Renda Familiar	Classe D - Até R\$ 600	39 (22%)
	Classe C - De R\$ 601 a 1200	56 (32%)
	Classe B - De R\$ 1201 a 2400	50 (29%)
	Classe A - Acima de R\$ 2401	30 (17%)